



1535

ARCHIVPFLEGE IN WESTFALEN UND LIPPE

Im Auftrage des Landschaftsverbandes Westfalen-Lippe
herausgegeben vom Westfälischen Archivamt, Warendorfer Straße 24, 4400 Münster

ZWEITES DEUTSCH-NIEDERLÄNDISCHES ARCHIVSYMPOSIUM
TWEDE NEDERLANDS-DUITS ARCHIEFSYMPOSIUM

in Arnheim

22. und 23. November 1981

PROGRAMM UND INHALTSVERZEICHNIS

DONNERSTAG, 22. OKTOBER 1981

10.30 Uhr	Empfang und Begrüßung der Teilnehmer	3
11.15 Uhr	Dr. H. Romeyk, Hauptstaatsarchiv Düsseldorf: <i>Zum Einsatz der EDV im Archivwesen</i>	3
12.15 Uhr	Diskussion	8
12.45 Uhr	Mittagessen auf Einladung der Stadt Bocholt	
14.00 Uhr	J. Roedoe, Min. Volkshuisv. en R.O., 's-Gravenhage: <i>The accessibility of archives and the computer today and in the future – the technical point of view</i>	9
14.45 Uhr	Diskussion	13
15.30 Uhr	Dr. J. Marschalck, Westfälisches Wirtschaftsarchiv, Dortmund: <i>Das Firmenarchiv Johann Caspar Harkort: Intensiverschließung mittels EDV</i>	16
15.50 Uhr	Dr. J.P. van de Voort, Visserijmuseum Vlaardingen: <i>Das MARDOC-Projekt</i>	19
16.10 Uhr	Dr. I. Auerbach, Institut für Archivwissenschaft, Marburg: <i>EDV-Arbeiten an der Archivschule Marburg und dem Hessischen Staatsarchiv Marburg in Zusammenarbeit mit dem Rechenzentrum der Universität Marburg</i>	23
16.30 Uhr	Diskussion	26
17.00 Uhr	Geselliges Zusammensein im Rijksarchief	
18.30 Uhr	Abendessen auf Einladung des Westfälischen Archivamtes	

FREITAG, 23. OKTOBER 1981

9.00 Uhr	Drs. P. Schotel, Gemeentearchief Dordrecht: <i>Increasing the accessibility of the notarial records of Dordrecht 1860 – 1869</i>	28
9.20 Uhr	H.P. Neuheuser, Archivberatungsstelle Rheinland, Köln: <i>Verzeichnung von Kommunalakten mit Hilfe der ADV</i>	29
9.40 Uhr	R. Riebschläger, Archivberatungsstelle Rheinland, Köln: <i>Übernahme und Erarbeitung von ADV-Programmen durch das Rechenzentrum des Landschaftsverbandes Rheinland</i>	34
10.00 Uhr	Diskussion	36
10.20 Uhr	Dr. M. Petry, Hauptstaatsarchiv Düsseldorf: <i>Erfahrungen beim Einsatz der EDV im Archiv</i>	38
10.40 Uhr	Drs. J. Spoor, Afd. Tekstsystemen Samsom Data Systemen: <i>Die Anwendung und Möglichkeiten der EDV bei der Erschließung von Archivgut</i>	43
11.10 Uhr	Schlußdiskussion	47
12.30 Uhr	Mittagessen auf Einladung der Provinz Gelderland	
15.00 Uhr	Besuch des Katholieke Documentatie Centrum (Kath. Universität Nijmegen) mit Vorführung des dortigen Computersystems Drs. P. Soetaert: <i>RAPIDE, Ein Informationssystem für die Erschließung verschiedenartiger Dokumentation</i>	50
	Verzeichnis der Teilnehmer	53

Vorbemerkung

Unter dem Thema *Erschließung von Archiven mittels EDV— Ontsluiting van archieven met behulp van de computer* fand am 22. und 23. November 1981 in Arnheim das zweite deutsch-niederländische Archivsymposium statt.¹

Die Organisation dieser Veranstaltung lag in Zusammenarbeit mit dem Stadtarchiv Bocholt in den Händen des für die Provinz Gelderland zuständigen Reichsarchivs Arnheim. Finanziell und materiell ermöglicht wurde die Tagung durch die Mitträgerschaft des Westfälischen Archivamtes, das auch die jetztige Publikation auf sich genommen hat, des Niederländischen Kultusministeriums (Ministerie van Cultuur, Recreatie en Maatschappelijk Werk) sowie der Provinz Gelderland, die dem Symposium im Haus der Provinz großzügig Unterkunft bot.

Um einen regen und fruchtbaren Gedankenaustausch zu gewährleisten, wurde nur eine beschränkte Zahl von Teilnehmern (etwa 25), überwiegend Archivare mit EDV-Erfahrung und Informatiker, geladen.

Im vorliegenden Heft finden Sie die Texte der an beiden Tagen vorgetragenen Referate sowie das Protokoll der sich anschließenden Diskussionen, die jeweils anfänglich Gelegenheit zu etwaigen Verständnisfragen boten und dann zur inhaltlichen Problematik übergingen. Das vor dieser Einleitung wiedergegebene Programm des Treffens mag also gleichsam als Inhaltsverzeichnis dieses Heftes dienen.

Die Organisatoren möchten dem Westfälischen Archivamt insbesondere dafür danken, daß es durch diese Publikation die Möglichkeit bietet, daß auch im größeren Kreise weitere Leute vom Fach noch nachträglich, wie man sagen könnte, am Symposium teilnehmen können.

Die Protokollführung wurde vom Kollegen P.F.T.M. Wouters, Reichsarchiv Gelderland, besorgt, der sich auf die unter Aufsicht von Herrn Sobeck, Stadt Bocholt, entstandenen Tonbandaufnahmen stützen konnte.

M. van Driel, H. Hofman
Rijksarchief Gelderland, Arnhem.
November 1981

¹ Das erste Archivsymposium wurde anlässlich der Internationalen Archivwochen im September 1979 in Bocholt veranstaltet. Das Tagungsprotokoll wurde in Heft 13, 1980 der Archivpflege in Westfalen und Lippe zum Druck gebracht.

DONNERSTAG, 22. OKTOBER 1981

Das Symposium wird eröffnet durch einige kurze Grußworte seitens der veranstaltenden Institute, ausgesprochen durch Stadtarchivar Dr. H. H. O p p e l für die Stadt Bocholt, Dr. H. R i c h t e r i n g, Ltd. Direktor des Westfälischen Archivamtes, und Dr. s. G. J. M e n t i n k, Rijksarchivaris in Gelderland.

Im Anschluß daran übernehmen Dr. F. D. J. K e t e l a a r und Dr. O p p e l die Moderation der Vormittags-sitzung. Als erster Referent erhält Dr. H. R o m e y k (Hauptstaatsarchiv Düsseldorf und Vorsitzender des EDV-Ausschusses des VdA) das Wort.

ZUM EINSATZ DER EDV IM ARCHIVWESEN

von Horst Romeyk

Der nachfolgende Bericht versteht sich nicht als *Soll und Haben der EDV* im Archiv, sondern als Aussage über die praktischen Erfahrungen aus den letzten 10 Jahren im Hauptstaatsarchiv Düsseldorf.

1. Die Notwendigkeit des EDV-Einsatzes

1.1 Intensivierung der Erschließung und Verzahnung der Findhilfsmittel

Die eigentliche Erschließung der Archivalien ist ihre Verzeichnung in klassifizierten, d.h. sachthematisch geordneten Findbüchern, die im Rahmen des Provenienzprinzips und über die Bestandsbezeichnung einen ersten Zugriff gestatten. Wenn auch mit Hilfe qualifizierter Beratung der Benutzer auf die wesentlichen, für sein Thema in Frage kommenden Bestände hingeführt wird, so gibt es doch vielfache Probleme, die sich nur durch bestandsübergreifende, aktualisierbare Findhilfsmittel bewältigen lassen. Diese Probleme ergeben sich vor allem aus den folgenden Sachverhalten:

In den meisten Archiven gibt es einen hohen Anteil an wachsenden Fonds, die häufig auf zahlreiche kleinere Findbücher verteilt sind; die Klassifikation führt zwar die gleichen Sachbetreffe zusammen, sie versagt aber bei personenbezogenen Forschungen oder lokalhistorischem Ansatz für die Ortsbetreffe. Darüber hinaus sind weitere Angaben in den Enthält- und Darin-Vermerken auch über eine vielfältig aufgeschlüsselte sachthematische Gliederung nicht erschließbar.

Neben die Vielzahl der Verwaltungszweige und betreuten Behörden tritt neu gegenüber dem herkömmlichen Selbstverständnis der Archive die Sammlung nichtstaatlichen Archivgutes, das gerade unter dem dieses Sammeln leitenden Gesichtspunkt der Ergänzung staatlicher Akten mit diesen durch entsprechende Hilfsmittel verzahnt werden muß. Auch durch die Intensivierung der Forschung, vor allem für solche thematischen Arbeiten, die mit örtlichen Quellen ggf. den Verlust der Akten

auf zentraler Ebene auszugleichen versuchen, werden erhöhte Anforderungen an die durchgängige Nachweisung der einschlägigen Betreffe gestellt. Nicht unterschätzen darf man schließlich, daß die Kollegien in den einzelnen Archiven heute größer sind und die Fluktuation stärker ist, so daß die Hilfsmittel unabhängig von der Person des jeweils zuständigen Bearbeiters greifbar sein müssen.

Aus dem letzteren Punkt allein schon ergibt sich weiter, daß auch die innerarchivischen Grenzen zwischen einzelnen Zuständigkeiten durch besondere Findhilfsmittel überwunden werden müssen.

1.2 Neue Findhilfsmittel

Als neue – oder neuartige – Findhilfsmittel wird man die bestandsübergreifenden Indices ansprechen können. Sie sind Zusammenspielungen der Einzelindices aus dem Bereich eines großen Registraturbildners oder die mehrerer paralleler gleichartiger Behörden.

Diese Hilfsmittel sind einzig über die EDV zu erstellen. Sie erreichen nämlich sehr bald Größenordnungen, die mit konventionellen Methoden nicht mehr zu bewältigen sind. Zettelkataloge von 20 – 30.000 Einzelkarten sind nicht mehr gut manipulierbar, sie sind auf jeden Fall ein Unikat. Einen Teil der Neuartigkeit darf man auch in der problemlosen Vervielfältigung sehen. Das Hauptstaatsarchiv Düsseldorf kann seine bestandsübergreifenden Indices sowohl den abliefernden Behörden als auch den übrigen staatlichen Archiven, soweit sie betroffen sind, zur Verfügung stellen und darüber hinaus auch im Benutzersaal in mehreren Exemplaren aufstellen.

1.3 Rationalisierung

Einer der Rationalisierungseffekte beim EDV-Einsatz liegt in der Entlastung im Bereich der Schreibearbeit. Gerade bei der Indexierung brauchen Angaben über die Fundstelle oft nur einmal für eine Vielzahl von Ein-

tragungen geschrieben zu werden. Die Verwendung von Zusätzen, die einzelne Eintragungen ergänzen, läßt sich auch über das Verarbeitungsprogramm steuern. Eine Zeitersparnis läßt sich nach verschiedenen Erfahrungswerten auf ca. 25% beziffern. – Sehr viel weitergehende Entlastung ist bei der Verwendung von Schreibautomaten oder auch Minicomputern für die Findbucherstellung zu erwarten.

Eine Rationalisierung ganz anderer Art ist über die Nutzung der EDV-Arbeiten in der öffentlichen Verwaltung denkbar. Die Nutzung integrativer Beziehungen zwischen den Verwaltungsbehörden und Archiven bietet sich vor allem da an, wo Namen und ansonsten formale Daten gespeichert sind. Es wäre theoretisch sicherlich möglich, bei den Behörden für die Zahlung der Besoldung und Versorgung die Personaldaten derjenigen Personen (in dem für das Archiv erforderlichen Umfang der Einzelangaben) zu sammeln, die aus dem öffentlichen Dienst ausscheiden und deren Personalakten mithin einmal an Archive gelangen. Dadurch könnte eine erhebliche Arbeitersparnis erzielt werden, da bei der Personalaktenarchivierung diese Angaben ja auch jeweils zu erfassen wären, und Fehlerquellen ausgeschaltet werden. Aus den Datensätzen ließen sich sicherlich auch Aufbewahrungskriterien ableiten. Einer solchen Verfahrensweise dürften die Bestimmungen des Datenschutzes nicht entgegenstehen.

Beispielsfall einer Zusammenarbeit zwischen Archiv und Behörde ist die Übernahme von Datensätzen zur Deutschen Grundkarte aus dem Landesvermessungsamt Nordrhein-Westfalen. Übermittelt wurden die Gauß-Krüger-Koordinaten, die Blattbezeichnung sowie die Nummer der Topographischen Karte 1:25 000. Als archivspezifische Angaben wurden in die Datei die Lagernummer des Archivs sowie Erscheinungsjahr und Blatttyp eingegeben. Die übermittelte Datei dient zusätzlich als Grundlage für eine Verzeichnung von Luftbildplänen, wobei auch wiederum nur einzelne Angaben nachgetragen werden müssen, um ein vollständiges Findhilfsmittel zu erhalten.

2. Leistungen der EDV aus archivarischer Sicht

Was sind nun die besonderen Leistungen, die die EDV für den Einsatz in den Archiven interessant machen? Muß nicht gewarnt werden vor einer vordergründigen Modernität des Einsatzes, weil es eben besonders fortschrittlich erscheint, ohne ernsthafte Prüfung der Berechtigung? Dies kann man nach langjähriger Erfahrung getrost verneinen. Es hat sich eine Reihe von Gründen klar herausgestellt, aus denen die EDV für die Archive attraktiv und notwendig ist.

Im einzelnen sind dies:

2.1 Aktualisierbarkeit

Daten in maschinenlesbarer Form lassen sich immer wieder korrigieren, ergänzen oder sonst fortschreiben und sind somit – freilich nach erneuter Verarbeitung per Programm – stets auf einem aktuellen Stand zu halten. Mit Hilfe der EDV lassen sich überdies Informationen, die meist unhandlich und nur einmal in Karteien gespeichert sind, in Form handlicher, vielfältiger und immer auf den neuesten Stand zu bringender Kataloge anbieten und in der gewünschten Zahl von Ausdrucken den Interessierten zur Verfügung stellen. Daraus ergibt sich eine sehr viel größere Flexibilität in den Findhilfsmitteln. Demgegenüber sind hand- oder maschinenschriftliche Indices auf ihre bestehende Form festgelegt: Korrekturen müssen manuell in allen Einzelexemplaren angebracht, ggf. diese sogar gänzlich neu geschrieben werden. Für das Ineinanderarbeiten mehrerer Indices müssen diese erst einmal völlig neu sortiert, anschließend wieder geschrieben und schließlich auch noch kollationiert werden.

Die Aktualisierbarkeit von Dateien öffnet noch weitere Möglichkeiten: Denkbar ist z.B. die Verwertung der Erkenntnisse, die bei der Recherchentätigkeit der Archive oder bei der Benutzung der Akten durch Außenstehende immer wieder gewonnen werden. Obwohl sie sicherlich allgemein von Wert für die weitere Erschließung der Archivalien wären, sind sie im Augenblick für die speziellen Interessen nicht von Belang. Dadurch aber, daß sie nicht in die Findhilfsmittel eingehen, bleiben sie auf den einzelnen beschränkt und geraten dann wieder in Vergessenheit. Für Bestände oder Bestandsgruppen mit aktualisierbaren Findhilfsmitteln könnten solche Informationen gesammelt, zusammengepielt und allgemein nutzbar gemacht werden.

Auch periodisch neu zu erstellende Übersichten – im Hauptstaatsarchiv Düsseldorf das Verzeichnis der Findbücher – oder auf die Dauer gar Veröffentlichungen (Beständeübersichten) können als Dateien leicht auf den neuesten Stand gebracht und wieder ausgedruckt werden.

2.2 Beliebige Reproduzierbarkeit gespeicherter Daten

Daten, die einmal in einer maschinenlesbaren Form gespeichert sind, können in immer neuen Zusammenhängen – freilich unter Berücksichtigung gewisser formaler Bedingungen – verwendet werden. Solche Anwendungsbereiche ergeben sich z.B. bei der Reper-torisierung, bei der Titelaufnahmen mehrfach ausgegeben werden können – auch in verschiedener Druckaufbereitung –, sei es innerhalb desselben Findbuchs, sei es in verschiedenen. Denkbar ist auch die Zusammenführung von gespeicherten Titelaufnahmen aus verschiedenen Beständen in einem sachthematischen Inventar. Sonder-

fälle der Reproduzierbarkeit sind die Einfügung vorab gegebener konstanter oder tabellarischer Werte zur Komplettierung von Angaben bei der Erfassung oder die nachträgliche stereotype Hinzufügung von Textteilen zur Erläuterung.

Die modernen Dienstleistungsprogramme, sei es bei der sog. Groß-EDV, sei es bei den Textautomaten, erlauben eine ganze Reihe arbeitsvereinfachender Prozeduren, die ebenfalls auf der Reproduzierbarkeit beruhen.

2.3 Verarbeitung erfaßter Daten nach verschiedenen Gesichtspunkten

Schließlich ist noch der Aspekt zu nennen, unter dem die EDV in die öffentliche Verwaltung eingeführt worden ist: die Verarbeitung großer Datenmengen, schnell und nach verschiedenen Gesichtspunkten.

Diese besonderen Leistungen der EDV, die zu ihrer allgemeinen und weitreichenden Anwendung geführt haben, gelten auch im Archivwesen, obwohl sie dort nicht unbedingt mit der gleichen Dringlichkeit für Problemlösungen eingesetzt werden müssen. Große Datenmengen werden aus der archivischen Verzeichnungs- und Erschließungstätigkeit selten auf einmal anfallen. Das Bedürfnis nach schneller Erledigung stellt sich auch nur bedingt, eher schon das der Verarbeitung nach verschiedenen Gesichtspunkten. Die Schnelligkeit und Präzision mit der vor allem Sortierungen durchgeführt werden, spricht aber doch dafür, solche Arbeiten eben nicht manuell zu erledigen – wobei überdies Fehler häufiger vorkommen, etwa beim Einsortieren von Nachträgen in Karteien –, sondern maschinell. Die Verarbeitung nach verschiedenen Gesichtspunkten ermöglicht in erster Linie, aus einmal erfaßten Daten ohne Mehraufwand anhand der gespeicherten Kriterien verschiedene Ergebnisse zu gewinnen.

3. Einsatzbereiche der EDV im Archiv

Die hauptsächlichlichen Anwendungsgebiete für die EDV im Archiv lassen sich nach den bisherigen Erfahrungen in drei Gruppen einteilen:

- Anwendung im Bereich von Übernahme und Bewertung von Schriftgut
- Anwendung zur Erschließung der Archivalien
- Anwendung als sog. Archivverwaltungsdateien.

Bei dieser Einteilung wird sicherlich der Schwerpunkt auf der EDV-Anwendung zur Erschließung der Archivalien liegen, während die beiden anderen Bereiche von der Größe des Archivs abhängen, in vielen Fällen sich aber durch Karteiführung genauso gut bewältigen lassen.

3.1 Übernahme und Bewertung von Schriftgut

Bis heute gibt es kaum allgemein anerkannte Richtlinien für die Bewertung des Schriftgutes, d.h. eine vorherige Festlegung der Archivwürdigkeit. Es wird auch mit Hilfe der EDV nicht zu Bewertungskriterien kommen, die etwa das Ergebnis eines "maschinellen Bewertungsprozesses" wären. Unzweifelhaft wird es aber über aktualisierbare Kataloge möglich sein, stärker als bisher alle praktischen Erfahrungen zu sammeln und fortzuschreiben.

3.2 Erschließung des Schriftgutes

3.2.1 Indexierung

Wie bereits ausgeführt, ist die Indexierung ein probates Mittel, die verstreuten Informationen so zusammenzuführen, daß dem Benutzer wesentliche Informationen nicht mangels Übersicht verlorengehen. Wichtig ist dabei, von vornherein die Verwendung eines Index für verschiedene Findhilfsmittel zu berücksichtigen. Im einzelnen handelt es sich darum, nicht nur die Seite des Findbuches, sondern auch dessen Signatur, die Bestandsbezeichnung sowie die Archivaliensignatur zu erfassen. Dies ist im Augenblick ein scheinbares Mehr an Arbeit, das sich aber auf jeden Fall auszahlt. Durch die Erfassung der Aktensignatur wird zugleich eine Konkordanz zwischen dieser und der Seitenangabe des Findbuches erstellt, die Erfassung von Findbuchsignatur und Bestandsbezeichnung ist nur einmal erforderlich und wird automatisch allen Einzeleintragungen zugeordnet.

Von ebenso großer Bedeutung ist die Unterteilung der Indices nach ihrem Inhalt. Es empfiehlt sich dringend, einzelne Teile getrennt zugreifbar zu machen, d.h. eventuelle Ortsindices, eventuelle Sachindices als Teildateien zu führen, die sich mühelos zusammenführen, aber nicht aus einer ungegliederten Datei herauslösen lassen.

Im Hauptstaatsarchiv Düsseldorf resultiert aus der getrennten Dateiführung:

1. der jeweilige Repertoriumsindex
2. der bestandsübergreifende Index auf der Grundlage von (1), ohne Personennennungen
3. der Generalpersonenindex, eine Sammlung aller Personenbetreffe (in einem abgegrenzten Bereich)
4. der sog. Generalindex, d.h. im wesentlichen die Betreffe aus dem Inhaltsverzeichnis der Findbücher.

3.2.2 Repertorisierung

Die Indexierung eines fertigen Findbuches bedeutet ungeachtet der in der EDV-Bearbeitung liegenden Ersparnisse in sich selbst noch eine ungeheure Doppelarbeit, muß doch praktisch ein gesamtes Findbuch in seinem Informationsgehalt noch einmal erfaßt, d.h. praktisch abgeschrieben werden. Wieviel sinnvoller wäre es, aus einem vorhandenen Text heraus die Indexbegriffe mehr oder minder automatisch zu gewinnen und ihren Fund-

stellen zuzuordnen? Findbücher für die neueren Bestände werden zudem in der Regel den Benutzern nur in einer maschinenschriftlichen Form vorgelegt. Es ist daher naheliegend, die ohnehin erforderliche Schreibarbeit gleich in maschinenlesbarer Form vorzunehmen.

Die archivische Titelaufnahme umfaßt sowohl formale (Datum, Laufzeit, Bandzahl usw.) als auch inhaltliche (Kurztitel, Enthält-Vermerke, Darin-Vermerke) Bestandteile, für die seit langem Erfassungsrichtlinien vorliegen. Diese einzelnen Elemente der Titelaufnahme lassen sich sehr wohl in verschieden strukturierte, in den je einzelnen Titelaufnahmen stets gleiche Einheiten fassen, für die als einheitliche Bezeichnung (in Anlehnung an im Bibliotheksbereich schon länger übliche Formen) "Kategorie" vorgeschlagen werden soll. Die Verwendung zweistelliger Codezahlen für die einzelnen Kategorien ergäbe insgesamt 100 mögliche (00 – 99) Elemente, womit eine ausreichende Untergliederung gegeben sein dürfte. Erforderlich wären Aufteilung und Reglementierung der einzelnen Kategorieninhalte.

Die Voraussetzung für eine maschinelle Verarbeitung ist die formale und inhaltliche Gleichartigkeit der einzelnen Bestandteile einer archivischen Titelaufnahme. Dazu gehören z.B. der gleiche Aufbau von Datumsangaben, gleichstrukturierte Signaturen (Bestandsbezeichnungen), einheitliche Gliederung in Kurztitel, Inhaltsanalyse (Enthält-Vermerke) sowie abweichende Dokumentationsinhalte. Unter diesen Voraussetzungen bietet die maschinelle Aufbereitung eines Findbuches keine besonderen Schwierigkeiten. Über das Verarbeitungsprogramm ist jedes gewünschte Druckbild steuerbar. Automatisch werden Inhaltsverzeichnis und Nummernweiser erstellt.

Unter dem Akronym "DRUIDE" (für "Düsseldorfer Repertorisierung und Indexierung durch EDV") wird zur Zeit ein System von Anweisungen und Programmen entwickelt, das bei der Einheit von Erfassung und Verarbeitung in einem Zuge das Repertorium und den Index (in seinen verschiedenen Teildateien für die weitergehende Verwendung) erstellen wird. – Für eine maschinelle Verarbeitung der Repertorien spricht auch die Möglichkeit der COM-Verfilmung (COM = Computer-Output-Microfilm), die sowohl den Bedürfnissen der Sicherungsverfilmung wie der Publikationsverfilmung entgegenkommt. Ebenso sind – für den letzteren Fall – Schnittstellen zum Fotosatz offen.

3.2.3 Konkordanzen und Übersichten

Die Erfassung des Gesamtrepertoriums unter dem Gesichtspunkt der maschinellen Weiterverarbeitung erübrigt eine Reihe zeitaufwendiger Anschlußarbeiten oder ermöglicht doch auch wieder die problemlose Erstellung von Hilfsmitteln, die vielleicht nicht unbedingt notwendig sind, aber die Arbeit im Bedarfsfall doch sehr erleichtern.

Dazu gehören zum einen Konkordanzen, die etwa den Ablauf von Sperrfristen erkennen lassen, indem man zu

der Aktensignatur das Endjahr der Laufzeit ausgibt. Vielfach muß zur Zeit noch umständlich nachgesucht werden, ob die Benutzung eines bestellten Archivals auch statthaft ist. Mit der Hilfe solcher Konkordanzprogramme lassen sich auch leichter Fehler aufspüren, d.h. eventuelle Doppel- oder Leernummern innerhalb eines Bestandes.

Übersichten ähnlicher Art sind außerdem solche über die kommenden Aktenzeichen. Für den Archivar sind sie ein Hilfsmittel beim Erkennen verwandter Überlieferungen, sie bieten Zuordnungen für die Klassifikation o.ä. Häufig aber sind – vor allem im Bereich der Gerichtsbarkeit – Aktenzeichen die Angaben, die bei den Archiven bei der Rückforderung für behördliche Zwecke gemacht werden.

3.3 Archivverwaltungsdateien

Abhängig von Größe und Benutzerverkehr eines Archivs sind einige ggf. wünschbare Dateien, mit deren Hilfe etwa der Archivalienversand, die Benutzungsgenehmigungen und eventuell die Benutzerthemen übersichtlicher gestaltet werden könnten. Besonders die beiden ersteren Punkte – d.h. die Kontrolle – bedeuten zeitaufwendige und lästige Arbeit.

4. Äußere und innere Erfordernisse und Voraussetzungen für den EDV-Einsatz

Zu den äußeren Voraussetzungen gehört in erster Linie der Zugang zu einer Datenverarbeitungsanlage. Sie erscheint, ungeachtet gewisser Schwierigkeiten, doch relativ unproblematisch. Der generelle Einsatz der EDV in fast allen Sparten der öffentlichen wie auch der Wirtschaftsverwaltung wird sicherlich eine Kapazitätsreserve für die – von der Rechnerbeanspruchung nicht allzu hohen – Bedürfnisse der Archive haben. Häufig sind aber nicht unerhebliche bürokratische Schwierigkeiten zu überwinden. – Zum anderen deutet die Entwicklung auf dem Gebiet der Schreibautomaten und auch der Mini-Computer auf eine recht verheißungsvolle Zukunft hin, sowohl von den Leistungen dieser Geräte als auch ihrem Preis her.

Als Grundstufe des Zugangs zur Datenverarbeitung ist vielleicht die Schreibmaschine mit Bildschirmerfassung (für das Repertorienschreiben oder als Datensammlersystem) anzusehen.

Eine erste Ausbaustufe wären mehrere Datenerfassungsplätze bzw. ein größerer "Textverarbeitungsautomat" mit den entsprechenden Speicher- und Verarbeitungsmöglichkeiten.

Sicherzustellen wäre der Anschluß an bestehende Großanlagen mit einem direkten Zugang über Terminal.

Ein gewisses Problem bleibt die Programmierung, die, sofern es sich um Sortiervorgänge handelt, zum großen Teil über Dienstprogramme (Systemsoftware) abgefangen wird. Auch für die Textverarbeitung und die damit

verbundene Indexierung stehen schon in erheblichem Umfang Standardprogramme zur Verfügung.

Für die Datenerfassung – sofern sie im Archiv selbst erfolgt – ist kein spezielles oder zusätzliches Personal vonnöten, da die Schreibarbeit, von gewissen äußeren Formalien abgesehen, den gewohnten Arbeiten im Archiv gleicht und sich durch die Erfassung in maschinenlesbarer Form sogar verringert (Verzicht auf besondere typographische Gestaltung, Vereinfachung der Korrekturen). Wichtig ist es aber, die Schreibkräfte mit den neuen Arbeitsweisen vertraut zu machen, um nicht unnötige Widerstände aufzubauen. Aus Unkenntnis über die Art der Erfassung entsteht häufig eine ablehnende Haltung, die in der Datenerfassung eine Art "Strafarbeit" sieht, die sonst niemand verrichten will.

Als Alternative zur Erfassung im Archiv bietet sich die Erfassung durch kommerzielle Unternehmen an. Damit sind allerdings zusätzliche Kosten verbunden. Gelegentlich haben auch größere Rechenzentren im Bereich der Verwaltung, in dem das Archiv angesiedelt ist, Möglichkeiten zur Datenerfassung.

Zu den inneren Voraussetzungen gehört in gewissem Rahmen eine Grundlagenarbeit, um den formalen Anforderungen des EDV-Einsatzes zu genügen. Dies sind:

- einheitliche und unkomplizierte Bestands- und Findbuchbezeichnungen. Hier sollten nach Möglichkeit numerische Signaturen ggf. auch Zahlen/Buchstaben-Kombinationen verwendet werden, aus denen bereits bestimmte Rückschlüsse auf die Tektonik des Archivs möglich sind.
- Verwendung einheitlicher Klassifikationen in gleichgelagerten Beständen ein- und desselben Registraturbildners, die in verschiedenen zeitlichen Tranchen abgegeben worden sind. Das gleiche gilt – mit Modifikationen – auch von Beständen, die von Behörden gleichen Typs abgeliefert werden. Für neuere Abgaben können (Muster-) Geschäftsverteilungspläne ein Anhalt sein.

Wichtig sind auch die vorgängigen Analysen der einzelnen Automationsvorhaben sowie die Dokumentation der Verfahrenslösungen. Der heilsame Zwang detaillierter Beschreibung erfordert eine gründliche Klärung der eigenen Vorstellungen, er führt zu der Zerlegung eines komplexen Vorganges in seine kleinsten Schritte und Einheiten und vereinfacht damit wesentlich die Programmierung. Die Analyse solcher Arbeitsschritte kann auch ohne spezielle Kenntnisse der Programmierung geleistet werden. Allerdings erweisen sich Kenntnisse für den mittelbaren Einsatz der EDV als sehr nützlich.

Klarheit muß schließlich herrschen über die langfristige Speicherung der erfaßten Daten. Es versteht sich von selbst, daß Texte, die von Zeit zu Zeit aktualisiert werden, auch in maschinenlesbarer Form verfügbar gehalten werden. Dasselbe gilt im Grunde von allen Dateien, da ja der große Vorzug ihrer Verwendung in anderen Zusammenhängen erst durch die maschinelle Verarbeitung gegeben ist. Alle maschinell verfügbaren

Dateien können in gewissem Umfang auch maschinell weiterbehandelt werden, z.B. die nur in Großschreibung vorhandenen Texte auf die gefälligere Groß- und Kleinschreibung umgestellt werden. Gespeicherte Texte eröffnen natürlich auch die Möglichkeit von Recherchen in den Dateien, doch wird im Bereich der nordrhein-westfälischen Staatsarchive nur für den begrenzten Zeitraum zwischen der Erfassung und vor der fertigen Verarbeitung ein solches Recherchieren ausnahmsweise zugelassen. Im übrigen ist es Prinzip, alle Ergebnisse aus Repertorisierung und Indexierung über die Listenausgabe den Benutzern zugänglich zu machen.

Abschließend sei bemerkt, daß der Arbeitsablauf beim Einsatz der EDV sich nicht grundsätzlich von den gewohnten Abläufen unterscheiden sollte. Dennoch hat die Erfahrung gezeigt, daß zum Teil große Verständnisschwierigkeiten auftreten, d.h. ob z.B. die von dem Verzeichnenden auszufüllenden Formblätter oder Belege usw. in die Maschine eingegeben werden, ob der Computer so etwas stets annimmt usw. Die allgemeine Schulung aller Mitarbeiter im Archiv sollte hierauf immer wieder eingehen. Voraussetzung für eine eingehende Schulung ist natürlich die einfache und unproblematische Handhabung der EDV-Anwendung. Wo sie den Charakter einer elitären Beschäftigung oder einer Geheimwissenschaft annimmt, wird sie sich nicht durchsetzen können.

D i s k u s s i o n

Nach dem Referat von Dr. R o m e y k erfolgte eine Diskussion, die einleitend Gelegenheit zu Verständnisfragen bot.

V a n d e V o o r t fragt, wieso der Vorgang der Klassifikation eine Erschließung durch die EDV beschränken könne.

R o m e y k: schränkt die Frage auf die Klassifikation von Serien ein und erläutert, die deutsche Archivwissenschaft verstehe unter Klassifizieren die systematische Ordnung der Archivtitel in Ober- und Untergruppen. Wenn eine Klassifikation eine Serie beinhalte, gebe ihre kurze Titelangabe keinen Zugang zu den in den Akten ferner vorhandenen Daten und Erkenntnissen. Was die verschiedenen Serienakten noch weiter an Personennamen und anderen Daten beinhalten, könne nur durch einen Index oder zusätzliche Titelaufnahmen erschlossen werden. Das aber bedeute unverhältnismäßig viel Arbeit.

M e n t i n k will wissen, wie man bei der Zusammenstellung einer Konkordanz (cf. Referat Punkt 3.2.3) Fehler im Inventar finde.

R o m e y k: Bei der Erschließung eines Aktenbestandes, nämlich der Aktentitelaufnahme und der Indizierung, werde eine Anzahl von Nummern und Signaturen gebraucht. Das Programm ermittle nachher, ob alle Nummern tatsächlich auch belegt seien oder etwa doppelt vorkämen. So erübrige sich eine sehr zeitaufwendige und anstrengende Kollationierungsarbeit.

F r e s e fragt, ob die moderne Vereinheitlichung von Orts- und Personennamen aus mittelalterlichen Urkunden bei der Indizierung nicht Identifikationsprobleme mit sich bringe.

R o m e y k verweist auf den Vortrag von Dr. Petry. Er glaubt, aufgrund seiner Erfahrung mit neueren Beständen, auch hier keine Schwierigkeiten sehen zu müssen. Freilich könne die eigentliche Identifikation von Orts- und Personennamen nicht durch die EDV abgenommen werden, sondern obliege auch weiterhin dem Bearbeiter des Urkundenbestandes.

O p p e l fragt, wie die Lücke zwischen den vom Benutzer gestellten Fragen und den im Computer eingespeicherten Daten überbrückt werde.

R o m e y k sieht die Lösung in einer guten Beratung des Benutzers und in der Anbringung von genügend häufigen Querverweisen.

K e t e l a a r fragt, ob für die Indizierung ein Thesaurus benutzt werde.

R o m e y k erwidert, die Erfahrung habe ihm gezeigt, daß die Ergebnisse seiner pragmatischen Arbeitsweise und eines Verfahrens unter Benutzung eines Thesaurus kaum Unterschiede aufwiesen.

O p p e l erbittet Erläuterungen über prozentuale oder absolute Platzersparnis angesichts umfangreicher Computeroutputs, wenn man statt dessen COM-fiches verwende.

R o m e y k bejaht dies in dem Sinne, daß es möglich sei, etwa 220 Ausdrücke (Output) auf einen Microfiche zu bringen. Außerdem bestehe die Möglichkeit, die auf Magnetbändern gespeicherten Daten auf Microfiche auszudrucken (COM). — Diese Frage führe allerdings zu weit von seinem Thema ab.

Die Frage K e t e l a a r s nach den Voraussetzungen und den Kosten des Einsatzes der EDV stellte R o m e y k hierbei bis zum Vortrag von Herrn Petry zurück. Man dürfe allerdings nicht glauben, daß schon der Einsatz eines Computers Anlaß zu direkten Personaleinsparungen gebe.

J a n s s e n erläutert ergänzend dazu, daß das Hauptstaatsarchiv Düsseldorf für die EDV-Arbeit eine Person des höheren und drei des gehobenen Dienstes ausgewiesen habe, die jedoch nicht alle ständig für diese Arbeit eingesetzt seien. Bei vier Personen bei einem Personalbestand von 130 Personen, könne daher nicht von einer finanziellen Mehrbelastung gesprochen werden, zumal die Resultate durchweg positiv seien. Das Hauptstaatsarchiv Düsseldorf sei auch ohnedies gezwungen, die EDV zu nutzen, schon allein um die vielen Neuzugänge zu bewältigen und darüber die Orientierung nicht zu verlieren.

K e t e l a a r resümiert, daß man in Holland noch zu wenig Erfahrung habe mit der EDV, weil die Verwaltungen ihre Archive erst nach einer Frist von 50 Jahren an die Archivdienste übergeben müßten und dürften. Dieser Rückstand gebe aber die Möglichkeit, von den Erfahrungen anderer zu lernen. Dies sei auch ein Zweck des Symposiums.

Nach dieser Diskussion fand auf Einladung der Stadt Bocholt das Mittagessen statt.

Am Nachmittag wurde die Sitzung unter Moderation von Dr. R o m e y k und Drs. C. O. A. S c h i m m e l p e n n i n c k v a n d e r O i j e fortgesetzt und mit einem Vortrag von Herrn J. R o e d o e eröffnet. (Managing Director des Rechenzentrums beim Ministerium für Wohnungswesen und Raumordnung (Min. van Volkshuisvesting en Ruimtelijke Ordening).

THE ACCESSIBILITY OF ARCHIVES AND THE COMPUTER TODAY AND IN THE FUTURE – THE TECHNICAL POINT OF VIEW

by J. Roedoe

You may wonder, why a managing director of the computer-centre of the Ministry of Housing and Physical Planning (V.R.O., Volkshuisvesting en Ruimtelijke Ordening) is delivering a lecture to an audience of archivists interested in computer applications. There is a simple explanation. Our computer-centre, when set up four years ago, was given the special task of doing research into office automation for the whole of the Dutch government administration. One of the aims of this work is the improvement of the accessibility of archives, both dynamic and static.

In the past, data-processing was especially based upon various kinds of structured data. The first data-processing was mainly card-based and was a complicated matter. We had to work with different types of punch-cards, in which there were columns, numbers, addresses etc. The first government computers in 1947 were running personnel salary-accounts—a typical example of structured data.

Over time, the possibilities for data-input developed. In the beginning, input went by way of cards punched on special machines. A card-reader would interpret these cards and make the input via a batch stream into the computer. Later on, some years ago, interactive workstations came into use. This meant, that one could type in parameters or other kinds of information and the machine would process these data immediately or respond to further questions. The workstation meant a revolution in dataprocessing.

Now, as far as the subject under consideration during our present meeting is concerned, where do the problems in data-processing start? Our problem is that we want to work with text containing unstructured data. A human being would structure the data easily by reading the text and underlining and/or placing key-words in a separate filing-system.

This is the way we would structure the data-items in a text; doing the same by means of a machine proved to be a little more difficult, so other methods were developed. One of these is to make use of a thesaurus.

Problems of a thesaurus

Such a system works like this. The machine reads the information (the unstructured text) and searches for the words of the text in the thesaurus. If a word is present in the thesaurus, it becomes a key-word. If not, it is considered irrelevant and stored in the machine as part of the text but not as a key-word for retrieval. This has been the first method of structuring information by computer-processing. In this system, the machine

would contain a list of key-words (the thesaurus) and in addition to this, an index of the numbers of the documents in which each of these key-words occurs. On the retrieval side, one would want to inspect the documents containing a specific key-word or combination of key-words. For example, I could use the key-word "man", putting it to the machine and finding, that it occurs in the documents numbered 105, 206, 2210. If I wanted to know something about the combination "man" and "machine", I might find that only one of these documents (2210) contains this combinations (see illustrations).

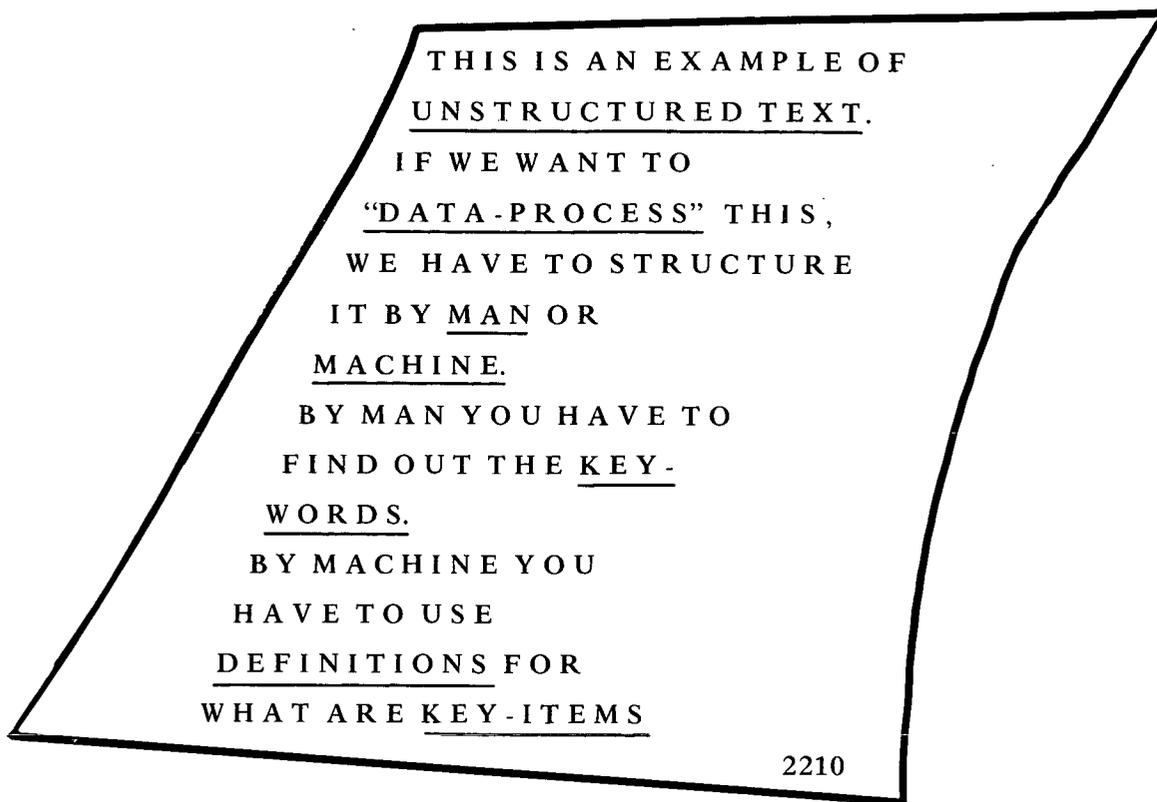
There are, however, some problems associated with this method. I think they will be known to you, as you are used to work in this way (although by hand, not by machine). One difficulty is the pre-defining of the key-words: what will be the key-words in the thesaurus? This is a very difficult problem, because one cannot know beforehand what variety of key-words will be in the system. For instance, proper names may be written in all kinds of ways, they may change over the years, and so on – this gives problems in predefining your thesaurus. Because of such difficulties, people have been searching for some other system of retrieval, in this case by machine and software.

One method is not to predefine the key-words, but on the contrary to define the stop-words, that is to say the non-relevant words in a text. For example, words like "and", "there", etc., may be defined as stop-words. When the machine reads text that is typed in, or inputted on cards or in whatever other way, it looks up all words in the text. If it encounters a stop-word, the machine decides that it is a non-relevant word and does nothing with it; all other words will be entered into a library of terms or key-words. This method of constructing a library or index of whatever you want to call it, is preferable because it is easier to predefine the stop-words and because all the terms that you might have forgotten to put into your thesaurus will now automatically become key-words for retrieval. In this manner a large variety of words will be key-words. Many software packages exist that can direct a computer to work like this, the packages "STATUS", "STAIRS", "DIRECT", "GOLEM" etc. Nearly every computer-manufacturer can supply a package that can do this kind of structuring.

Handwritten documents and images

Your problems really start, if you have work with unstructured data such as hand-written documents, or with a combination of images and text. Then you are in trouble. At present, the machine cannot read these kinds of information. You do not only have to work

UNSTRUCTURED TEXT



STRUCTURED BY MAN

KEYWORD:	DOCUMENT:
UNSTRUCTURED TEXT	1021, 503, 2210
TEXT	1021, 503, 606, 2210
UNSTRUCTURED	1021, 503, 708, 2210
DATA-PROCESS	541, 2210
MAN	105, 206, 2210
MACHINE	125, 2210
KEYWORDS	203, 503, 2210
DEFINITIONS	123, 205, 10, 2210
KEY ITEMS	12, 1038, 2210

STRUCTURED BY MACHINE

with text-items, you also have to work with images and (and I think this is very important) there will be relationships between text and images. There will be images which are part of a text, text which is part of an image, and you will want to input this relationship into the machine and eventually you will want to search and retrieve on image or data or the possible combinations. For example, if I wanted to have a picture of Jan, I would type "Jan" into the machine and it would have to provide me with the image. If I wanted to find a person with blonde hair the computer would have to be able to produce the image according to the description in the text. Thus, the relation between text and images presents a problem.

Image-Processing and its methods

A new medium has been developed for this kind of application, called "IMAGE PROCESSING".

Image processing works with a graphic image. These images are worked with directly on a computer screen. Imagine that I am looking at data information – an image – on the screen (how the image was put into the machine, I will explain later), and a part of this image shows a mountain. If I want to know more about this mountain, I can indicate this to the computer and it will provide me with a scale image of the mountain, or with statistics about the mountain or with whatever other information is available. The same can be done with all kinds of hand-written information. Suppose, I have someone's name written in a very old letter, and I want to find this name in all my other documents, then I can store this name as it is written and the machine will recognize it in all other documents and provide me with all relevant documents. This is what we call image processing i.e. direct processing with the graphic image of certain data.

In fact the computer can compare two images and if they are the same or nearly the same, he can indicate this, and show it to me.

Within image-processing there are 4 different steps:

The first step is writing the image into an electronic file. By electronic file I mean any file that can be written and read by the machine – this is the first problem.

The second step is image-handling, which, as I have already explained, is the ability to produce written information about (part of) an image or to check if it recurs in other images.

The third step is the relationship between image and data which must be defined and stored.

The fourth step is image-output. If I have processed the image, and I want to print it in a book, or in general on paper, I must be able to create output from the electronic file.

As for the first step (how to put the picture into an

electronic file), there are different ways of doing this, e.g. by using a television camera or a video camera. I can put my image under the camera and take a photo of it. This can be done with the same resolution as is now used by T.V. (1625 lines to an image), but it may also be done with a higher resolution. The camera then relays the video image to the tube or screen, making it visible in the same way as your T.V.-set at home does. Next, some type of processor digitizes the image, which means that it assesses it for any differences in colour, differences in strings etc. The image is made "processable" by a digital information system. Finally, the image is put on a storage medium, which may be a magnetic tape, a magnetic disc or some other type of medium. This method of image-filing is often used and can produce just as high a quality as conventional photography does. Another way of electronics filing of documents is by feeding the document to a machine which can read it and output it in a form similar to a photocopy, which is immediately put into a magnetic disc or some other storage medium. This is an alternative method which presents all sorts of difficulties because the original document must pass through the machine, and I am always in doubt, as to whether the document will ever reappear!

Another problem if you want to have good quality images is, that you will need a resolution of about 100 dots per mm². If we are talking about A4 format, there will be about 6 million dots per image which represents about the same amount of computer bits (electronical dots). Translated into bytes or characters, this means about 1 million characters on the storage medium. This is the same amount as needed for 250 pages of text, so you will understand, that we have a storage problem here. Of course I can use a lower resolution needing less storage capacity but this would be detrimental to the quality of the image.

The D.O.R.-system – Capacity and costs

This is one of the biggest problems of the magnetic media used until now. The magnetic disc has only a 200 megabyte capacity i.e. 200 pictures. In the case of an exchangeable disc or an equivalent fixed disc, the capacity becomes 600 images. What are the costs of this image filing? The disc including unit costs about 100.000 guilders. If exchangeable discs are used the costs are 2.000 guilders per disc (excluding unit).

In practice, however, it is inefficient to have to change a disc every time some other information is required. This is the reason why image processing in this way and with this kind of storage media is not practical. It is too expensive in comparison with any manual system.

However, a solution to the problem is currently being developed: "Digital Optical Recording" (D.O.R.). On this kind of medium an incredible amount of information can be stored and this information will be static.

This technology has been developed by PHILIPS in the Netherlands, but D.O.R. will in future be manufactured by nearly every computer manufacturer. At this moment IBM is busy working on it, as are ICL, Control Data, to name but a few. One problem is that PHILIPS often develops a new technology, but that production is impossible because our country is too small for sufficient investment to start the production of a product like this. For that reason PHILIPS sold the technology to other computer manufacturers. Some of these will have working D.O.R.'s available about 1983.

Some facts and figures on D.O.R. The diameter of the disc used in D.O.R. is about 30 cm, the same as a normal gramophone record. The storage capacity and this is very important, is about 10.000 M per disc, so you can put about 10.000 images on one disc. The price of the disc is at this moment about 400 guilders, but in the future it will drop to about 100 guilders per disc. Another interesting aspect is the price of the read-only unit, which will be about 10.000 guilders in the future. Currently they can be purchased in the States for about \$ 7.000, so if it is produced here, in Europe, it will cost about 10.000 guilders. The D.O.R. is static, which means that information written on it cannot be erased. It is certain at this moment that the information will remain on the disc for about 50 years without loss of quality. A magnetic media in comparison must be "refreshed" every 6 months. D.O.R. is only of interest if you are talking in terms of large amounts of information. You must understand that if I store written information on a disc like this, I can store the 500.000 pages of the "Encyclopaedia Britannica" on one disc.

Possibilities of using computers in the archives

How can you as archivists work with these implements in the future? What you must do first is put the image on the disc. This would be the first step. You could do it yourself but I think that all the necessary processing required for this one-time creation of images might just as well be done in batch mode by a service bureau. Computerservice centres or universities will have all the necessary facilities. You could tell the computer-centre: here are 10.000 pages from our archives, here is my disc, please store the material on the disc.

Next, you would have to define for the computer the relationship within the data i.e. the data-elements which belong to a specific image. You could do that on a micro-computer, an "Apple" or something similar, but if you'd want to use the interesting packages such as "STATUS", "STAIRS", "GOLEM" etc. you might do it using a terminal linked to a big computer system. You would be using only a very small part of a computer like that, so in that case it would be advisable to rent storage space on a large mainframe.

Now, if you wanted to start processing your images, what would you have to do? You'd submit your selection parameters to the data-computer, or the micro-computer-system, and the computer system would supply you

with the addresses of the relevant images that are on the D.O.R. The read-only D.O.R.-unit would go to your address and provide you with an image with which to work. The handling of the image, the image-processing itself would then take place in the terminal as a part of the terminal hardware. According to your wishes, the handling might also take place elsewhere. The workstations used for this sort of processing are still a little expensive. On the screen you see the picture consisting of terms or images or whatever, and you can indicate, using a pen or other media, which part of the text you want to structure, enlarge or replace etc.

At this moment workstations like this cost between 40.000 and 80.000 Dutch guilders, in future they will possibly cost about 20.000 guilders.

I have already told you a little about practical considerations. On site you can work with a read-only D.O.R. unit, which quite resembles a gramophone. You place the disc on it, you give the address and the unit finds the image for you and displays it on the screen. At this point you can do everything with the text, print it, play with it etc.

For hardware, you would need a micro-processor for supplying you with the addresses, but you could also use another data-processing system somewhere else, if you wanted to use the big packages. You would also need an image-processor on site, that is to say an intelligent screen. The data-processing itself may take place on your own premises or elsewhere linked by a terminal – that is irrelevant.

The price of the read-only unit will be very low as it will be on sale in the supermarket. The video-L.P. is already in use and since this is the same medium, it will therefore be massproduced and the price will be low; the price of the disc is \pm 100 guilders, of the reading unit in future perhaps 2.000 or 3.000 guilders.

CAFS-system

This is what I wanted to tell you about future developments. I think, that every data-processing system is based upon the fact, that we have limited facilities for storing our images. This is a big problem. It is currently too expensive to have big files for storing this kind of information, but with the arrival of D.O.R. it will be possible to store information very cheaply. The price of storing information on D.O.R. is less than the price of conventional storage on paper, and the price is still dropping.

There are some other developments, but their practicability depends extensively on the organization of the data-processing. In England Content Addressable File Storage (CAFS) has been developed. This is a system that can search on incomplete key-words. You need only define one letter or a few letters of the key-word and CAFS can search for it very quickly. This is, however, a machine that you can only use in a large centralized organization, because the machine costs about

1 million guilders. On the other hand, hundreds of terminals can use it simultaneously, so you will never find or need a machine like that in just one archive-building.

CAFS will in the future provide a solution for centralized archive-administrations and this is the application intended for this kind of product. However, this always poses problems because, as we have seen in the past, if you attempt coordinated development and use of systems, you will find that all partners will want to develop their own systems for their own use.

I think in archiving this will also create problems, you'll want to develop your own products and realize your own ideas. This is why I think, that a product like CAFS will not be popular in this application environment. Only processing based upon the real images can solve your problems; with D.O.R., in 2 or 3 years this will come out of the laboratory-stage into production and will be used on a large scale.

Diskussion

K e t e l a r meint, es sei noch ein Problem, ein Papier des Formats DIN A 4 auf dem Bildschirm deutlich sichtbar und lesbar zu machen.

R o e d o e antwortet darauf, Xerox habe inzwischen einen Bildschirm entwickelt, der größer sei als ein DIN A 4-Blatt, auf dem somit noch zusätzliche Informationen gegeben werden könnten. Die Wiedergabe [ndl.: oplossend vermogen] auf einem Bildschirm sei bis vor kurzer Zeit zu grob gewesen, aber auch hier habe die Technik Abhilfe geschaffen. Es seien Bildschirme in Entwicklung, die eine Informationsdichte von 100 dots/mm² aufwiesen, was durch konventionelle graphische Verfahren nicht erreicht werden könne.

Ein anderes Beispiel seien die Luftfotos, die für das Kataster in den Niederlanden gemacht worden sind. 8000 Fotos dafür seien aus 6000 Meter Höhe aufgenommen worden. Bei einer Wiedergabe auf dem Bildschirm könnten noch Objekte von 20 cm sichtbar gemacht werden.

K e t e l a r fragt nach den Kosten des Digital Optical Recording (D.O.R.).

R o e d o e schätzt einen Recording-unit inclusive Processor auf ungefähr 100.000 DM. Ein Terminal, der an einer zentralen Einheit angeschlossen sei und die Möglichkeit des Lesens und Schreibens biete, werde ungefähr 10.000 DM kosten. Gefragt wird ebenfalls, ob die Bilder auch schon in Farbe abgedruckt werden können.

R o e d o e bestätigt dies, fügt aber hinzu, daß eine derartige Anlage ungefähr 400.000 DM kosten dürfte. Er glaube allerdings nicht, daß sich Farbdrucke für ein Archiv lohnten.

S c h o t e l fragt, ob das D.O.R.-System mit den normalen Microfiches konkurrieren könne.

R o e d o e bejaht dies; D.O.R. habe auch den Vorteil, daß der Arbeitsverlauf vollständig elektronisch sei.

S c h o t e l fragt weiter, ob es sich lohne, mit Mikrofilmprojekten zu warten, bis das D.O.R.-System ausgereift sei.

R o e d o e meint, es sei schwierig, hier eine Prognose zu treffen, doch habe das System eine sichere Zukunft. Das Ministerium für Volkshuisvesting en Ruimtelijke Ordening habe jedenfalls die Microverfilmung bereits eingestellt.

S c h o t e l fragt, ob es eine Begrenzung in den Formaten der Originale geben werde.

R o e d o e sieht hier keine Schwierigkeiten, weil alles, was man mit einer Videokamera aufnehme, auch auf den Bildschirm projiziert werden könne.

R i e b s c h l ä g e r erkundigt sich nach den Zugangsmöglichkeiten zu den gespeicherten Informationen.

R o e d o e erklärt dazu, daß erstens Daten in den Computer eingegeben und diese mittels eines Index in den eingespeicherten Akten wiedergefunden werden können. Zweitens kann mittels eines Programms auch ein bestimmtes Bild in den Computer eingegeben werden, das die Maschine wiedererkennen und zurückfinden kann.

K e t e l a a r schließt hieraus, daß man also nicht an bestimmte festgelegte Bilder gebunden sei, sondern daß es dann auch möglich sei, handgeschriebene Texte, Worte und Buchstaben wiederaufzufinden.

R o e d o e bestätigt dies unter der Bedingung, daß die Maschine ausreichend programmiert sei, *was von den Anwesenden mit Skepsis aufgenommen wurde.*

V a n D r i e l fragt, wie der Computer einen Namen, z.B. Roedoe, auffinden kann, wenn dieser von 30 verschiedenen Personen geschrieben sei, und wies damit auf die Probleme der paläographischen Variationen hin.

R o e d o e: In Japan gebe es bereits Computer, die fähig seien, handschriftliche Postleitzahlen zu lesen und zu sortieren. Es gebe zwar noch Schwierigkeiten und der zeitliche wie finanzielle Aufwand seien sehr groß, es sei aber prinzipiell möglich!

S o e t a e r t fragt, was die Kanadier in diesem Bereich geleistet hätten.

K e t e l a a r verweist auf das ADPA-Bulletin des Automationsausschusses des International Council on Archives hin, und fügt hinzu, daß deren Ergebnisse sehr erfolversprechend seien.

V a n d e V o o r t weist auf die guten Resultate in Frankreich hin.

R o e d o e bestätigt dies und bemerkt, Frankreich stehe in diesem Forschungszweig ganz vorne. Aber alle Systeme – auch Philips – befänden sich noch in der Laborphase.

O p p e l erbittet Informationen über die Haltbarkeit der Bildplatten, die angeblich 50 Jahre haltbar seien.

R o e d o e erwähnt Laborversuche, die eine Haltbarkeit von mindestens 50 Jahren ohne Informationsverluste zeitig hätten. Nach Ablauf dieser Frist sei ein Überspielen der Platten auf neue möglich, was derzeit mit einem Kostenaufwand von 100,- DM je Platte verbunden sei.

S o e t a e r t macht auf das Problem der Änderung und Verbesserung der eingegebenen Daten aufmerksam, die bei Bildplatten nicht möglich seien.

R o e d o e sieht hier weniger Probleme, weil es um statische Bestände gehe, die sich nicht änderten. Man könne aber einen Teil der Bildplatte für Ergänzungen und Verbesserungen reservieren oder diese beim Überspielen der Platte einbringen.

J a n s s e n äußert, daß es sich in der Diskussion am Vormittag um die E r s c h l i e ß u n g von Archiven mittels EDV gehandelt habe, beim D.O.R.-System gehe es aber eigentlich mehr um den E r s a t z von Originalarchivbeständen. Er frage sich deshalb, ob die Erschließung mittels Bildplatte nicht arbeitsaufwendiger und teurer würde als direkte Erschließung der Bestände.

R o e d o e erklärt, da sei kein Unterschied zu erkennen. Die Eingabe und Verarbeitung sei dem heutigen Stand anderer Methoden gleichzustellen. Die gesuchte Information werde sogar schneller gefunden, könne gleich kopiert und herausgegeben werden. Außerdem habe D.O.R. den Vorteil, daß die Bildplatten stets dort gelagert werden könnten, wo sie gebraucht werden.

K e t e l a a r schließt sich der Meinung von Prof. Janssen an und betont, daß die Erschließung beim D.O.R. vorher erfolgen müsse.

J a n s s e n fügt hinzu, es gebe viele Archibände oder Liassen, die weder numeriert noch sonstwie gekennzeichnet seien. Dies müsse aber vor der Speicherung mittels D.O.R. geschehen und verursache allein schon einen großen Kostenaufwand.

S c h o o n d e r b e e k unterstreicht zusätzlich noch die hohen Kosten eines Computers, der lesen und schreiben könne.

R o e d o e äußert die Meinung, daß die Kosten von ungefähr 100.000 DM geringer als die Kosten für zwei Personen seien. Angesichts eines Personalbestandes von 130 Personen sei dieser Posten kein erheblicher Faktor.

S c h o o n d e r b e e k bemerkt dazu, er kenne in Holland noch kein Archiv, das sich auch nur einen einfachen Computer leisten und bezahlen könne.

Nach R o e d o e biete D.O.R. aber so viele Anwendungsmöglichkeiten, daß es vielseitig verwandt und somit zur Kostensenkung eingesetzt werden könne.

J a n s s e n fragt, ob es möglich sei, während der Speicherung neue Fragestellungen hinzuzufügen, die anfangs nicht angegeben seien.

R o e d o e betont, die Lösung des Problems liege im

Thesaurus und/oder geradewegs in der Anpassung des Programmes.

R o m e y k faßt dahin zusammen, daß die Technik mehr Möglichkeiten biete als der Mensch im Moment für möglich halte. Man tue gut daran, sich vorher über Möglichkeiten und Kosten zu erkundigen, etwa in Gesprächen mit Fachkollegen, um aus deren Erfahrungen und Wissensstand zu lernen, was dieses Symposium ja auch beabsichtige.

Romeyk fordert daher die Kollegen auf, ihre eigenen Erfahrungen mit EDV insbesondere unter dem Gesichtspunkt der Kosten und des Nutzens mitzuteilen und darüber in eine Diskussion zu treten.

S c h o t e l sieht für ein kleines Archiv wie das Stadtarchiv Dordrecht überhaupt keine Möglichkeit, sich einen Computer zu leisten. Mit der Erasmus-Universität Rotterdam sei daher vereinbart worden, daß sie den Computer stellt sowie die Hard- und Software liefert, das Archiv hingegen den Arbeitsaufwand erbringe.

Frau A u e r b a c h macht einen Kosten-Nutzen-Vergleich der HETRINA- und HESAUS-Projekte in Marburg. Monatlich kommen ca. 15 Anfragen von Amerikanern, die hieraus Auskunft haben möchten. Für diese Informationen bezahlen sie DM 16,—. Der gehabte Kostenaufwand sei aber vergleichbar mit acht Jahresgehältern eines Archivberrates, also verzinse die Sache sich kaum.

R o m e y k sieht in der Entwicklung und Verbreitung der Minicomputer die Möglichkeit für jedes Archiv, sich einen solchen zu beschaffen. Außerdem könnte eine Zusammenarbeit mit anderen Institutionen wie Universitäten oder Verwaltungen erwirkt werden, deren größere Computer oft nicht voll ausgelastet seien und somit vom Archiv mitgenutzt werden könnten. Herr Schotel habe bereits auf das Beispiel einer Kooperation zwischen einem Stadtarchiv und der Universität hingewiesen.

R o e d o e meint, daß eine Kosten-Nutzen-Analyse auf jeden Fall angestellt werden müsse. In seinem Ministerium z.B. mußten täglich 400 Telefonate bezüglich Beihilfeanträge bearbeitet werden, und zwar alle durch herkömmliche Recherchiermethoden. Folge: ungefähr 10.000 Gulden im Monat weniger an Telefonkosten.

S c h i m m e l p e n n i n c k fragt, wieviel Nachfrage vorhanden sein müsse, um die Kosten zu bestreiten und führt aus, daß ein Archiv bei weitem nicht soviel Anfragen zu bearbeiten habe wie eine amtliche Stelle, die ihre Altregistratur 50 Jahre bis zur Abgabe ins Endarchiv selbst betreue.

R o e d o e erwiderte, daß der Wert eines Archivbestandes in erheblichem Maße abhängig sei von der Zugänglichkeit, welche eben die Benutzung fördere.

R o m e y k bezweifelt dennoch, daß die wissenschaftliche Forschung sich in diesem Maße ausdehnen lasse und der Einsatz von Computern unter diesem Gesichtspunkt rentabel werde.

Nach dieser Diskussion folgten die Referate von D r . J . M a r s c h a l c k (Westfälisches Wirtschaftsarchiv Dortmund), D r s . J . P . v a n d e V o o r t (MARDOC-Project, Fischereimuseum Vlaarding) und D r . A u e r b a c h (Archivschule Marburg).

DAS FIRMENARCHIV JOHANN CASPAR HARKORT: INTENSIVERSCHLIESSUNG MITTELS EDV

von Juliane Marschalck

Im Folgenden soll berichtet werden über ein Projekt zur Intensiverschließung eines Firmenarchivs, das im März 1979 im Westfälischen Wirtschaftsarchiv in Dortmund begonnen wurde und das am 30.9.1981 zu einem vorläufigen Ende gekommen ist. Im Rahmen dieses Projekts wurde auch mit dem Einsatz der EDV für eine solche Aufgabe experimentiert.

Gegenstand des Erschließungsvorhabens war das Archiv der Firma Johann Caspar Harkort, das sich seit April 1978 als Depositum im Westfälischen Wirtschaftsarchiv befindet. Die Familie Harkort, seit der Mitte des 16. Jahrhunderts auf dem Gut Harkorten bei Hagen nachweisbar, betrieb Rohstahl- und Sensenhämmer und exportierte Eisenwaren nach Nord-, Nordostdeutschland und nach Skandinavien. In den 30er Jahren des 19. Jahrhunderts vollzog sich die Umwandlung der Handelsfirma in eine Fabrik für Maschinen, Maschinenteile und Eisenbahnbedarf. Ab 1860 wurde auch Eisenkonstruktions- und Brückenbau betrieben. Heute ist die Firma erloschen.

Gründe für die Auswahl gerade dieses Bestandes für eine Intensiverschließung waren:

1. die Repräsentativität der Firma für den im bergisch-märkischen Raum häufig anzutreffenden Typ des Familienunternehmens mittlerer Größe, mit weit zurückreichender Tradition und in einem für die Grafschaft Mark ebenfalls typischen Gewerbebereich, der Eisenverarbeitung;
2. der Umfang und relative Geschlossenheit des Bestandes: es wurde zunächst ausgegangen von 123 Geschäftsbüchern und rund 54.000 Geschäftsbriefen der Empfängerkorrespondenz aus der Zeit von 1712 – 1896. Eine Revision der Bücher gelegentlich der Vorbereitungen zur Sicherungsverfilmung ergab dann sogar eine Anzahl von 167 Geschäftsbüchern. Bezieht man die ab 1750 in Kopiebüchern vorliegenden Briefkopien mit ein, so erhöht sich die Anzahl der erhaltenen Geschäftsbriefe auf 70 – 80.000.
3. die Gefährdung des Bestandes wegen des teilweise sehr schlechten Erhaltungszustandes von Büchern und Briefen.

Von Anfang an war das Projekt stark an den Bedürfnissen potentieller Benutzer aus dem Bereich der Sozial- und Wirtschaftsgeschichte orientiert, weshalb wiederholt Kontakte mit Vertretern der einschlägigen Lehrstühle der Universitäten Bochum und Bielefeld gesucht wurden. Gedacht war nicht nur an die Erstellung eines üblichen Findmittels, sondern auch daran, eine bisher wenig benutzte, weil schwer zugängliche Quellengattung besser aufzuschließen. Ziel war zunächst die Erschließung des Bestands durch Personen- und Ortsregister, die auch im Druck publiziert werden sollten. In

diesem Sinne wurde der Antrag an die VW-Stiftung gestellt, die das Vorhaben finanziert hat. Denn daß eine derartige Arbeit im Rahmen der normalen Archivroutine nicht zu bewältigen ist, dürfte einleuchten.

Formale und inhaltliche Struktur des Bestandes

Die vorhandene Überlieferung läßt sich gliedern in

1. Hauptbücher
2. Briefkopiebücher
3. Memoriale
4. sonstige Nebenbücher
5. Empfängerkorrespondenz.

Diese Gliederung wurde auch bei der Signaturvergabe zugrunde gelegt. Während die unter 1. bis 3. genannten Bücher jeweils in längeren Serien vorliegen, finden sich von den kleinen Nebenbüchern stets nur einzelne Exemplare.

Die Überlieferungsdichte ist im Zeitraum von 1800 – 1850 am größten, d.h. gerade die wirtschafts- und sozialgeschichtlich hochinteressante Phase des Übergangs in die Frühindustrialisierung – im Fall Harkort besonders reizvoll wegen der engen geschäftlichen Beziehungen der Brüder Johann Caspar (V.), Friedrich und Carl & Gustav Harkort – ist außerordentlich gut dokumentiert.

Die inhaltliche Sichtung des Materials ergab eine wesentlich größere Reichhaltigkeit und Vielfalt an Informationen als erwartet. Das gilt vornehmlich für die Korrespondenz, aber auch für die Memoriale bis mindestens ins erste Drittel des 19. Jahrhunderts. Die Hauptbücher gestatten zwar einen Einblick in den Kundenkreis und den Gesamtumsatz, erlauben indes den Zugriff auf einzelne Geschäftsvorgänge und Details des Warensortiments allenfalls im 18. Jahrhundert.

Bearbeitung und Erfassung

Aufgrund der bisher skizzierten Feststellungen wurde ein über den ursprünglichen Ansatz hinausgehendes Bearbeitungskonzept entworfen. Ausgehend von der Korrespondenz als Rückgratserie des Bestandes wurde vorgesehen, dem Personen- und Ortsregister auch einen Sachindex hinzuzufügen. Dazu empfahl es sich, die Briefe einzeln zu verzeichnen. Angesichts der unter solchen Umständen zu erwartenden Datenfülle bot sich der Einsatz von EDV an. Hier schienen sich Möglichkeiten sowohl für eine automatische Registraturerstellung zu ergeben als auch für eine eventuelle Verfügbarmachung der auf Magnetband gespeicherten Daten für Benutzer-

recherchen. Es galt daher, einen Erfassungsmodus zu finden, der die spätere Anwendung von EDV zumindest offenhielt. Grundsätze für die Erfassung mußten demnach sein:

- ein Mindestmaß an Formalisierung der ausgewählten Informationen,
- eine möglichst umfangreiche Verzeichnung von Details, um das aufbereitete Material auch für zukünftige Forschungsfragen möglichst offen zu halten,
- Verzicht auf die vorgängige Erstellung eines Thesaurus für den Index.

Da es im deutschsprachigen archivischen Bereich kein Vorbild für eine derartige Verzeichnung gibt, mußte ein geeigneter Erfassungsbogen selbst entwickelt werden. Der besseren Handhabbarkeit wegen wurde das Format einer DIN-A-5-Karteikarte gewählt. Sie enthält im "Kopf" formale Angaben zum jeweiligen Brief: Signatur, Name der Firma oder Person des Korrespondenten, Sitz der Firma bzw. Wohnort der Person, Datum des Briefs, Angaben zu Beruf oder Gewerbszweig des Korrespondenten sowie eine Identifikation des Briefs als ein- oder ausgegangen. Personen- bzw. Firmennamen werden in zwei Versionen aufgenommen: in einer normierten Form, nach der alphabetisiert wird, und in der jeweiligen Originalschreibweise. Diese Einrichtung erwies sich als notwendig wegen der orthographischen Unsicherheiten des 18. und frühen 19. Jahrhunderts.

Eine Rubrik "inhaltliche Erschließung" soll – unformatiert, knapp und in möglichst standardisierter Terminologie – einen Überblick über den Briefinhalt geben: Charakterisierung des/der betreffenden Geschäftsvorgänge (Bestellung, Rechnung, Reklamation etc.), Warenart(en), falls erwähnt, Transportwege, Wechselplätze, sonstige geschäftliche und private Nachrichten (z.B. Informationen über die Solvenz von Kunden, über Konkurse, Verkaufsstrategien etc.). Es handelt sich hierbei also nicht um ein Regest.

Eine Spalte "Bemerkungen" ist vorgesehen für Betreffende, die sich nur schwer bei der Inhaltsangabe berücksichtigen lassen (etwa Informationen über Verwandtschaftsverhältnisse der Korrespondenten, Abfassung des Briefs in einer Fremdsprache u.ä.).

Schließlich werden auf der Rückseite der Karte die Namen aller Personen und Firmen aufgelistet, die in dem betreffenden Brief genannt werden. Viele Kunden – wie die Landkrämer und "Kesselführer" in Norddeutschland – und auch Lieferanten der Firma sind nur auf diesem Wege zu erfassen, da sie als Korrespondenten nicht in Erscheinung treten, sei es, daß die Überlieferung verlorengegangen ist, sei es, daß eine solche gar nicht existiert hat.

Parallel zur Aufnahme der Einzelverzeichnung der Briefe wurde eine Reihe von Informations- und Beratungsgesprächen mit EDV-Benutzern aus dem archivischen Bereich sowie mit DV-Fachleuten geführt, um konkrete Möglichkeiten für die elektronische Bearbeitung der aufbereiteten Daten zu eröffnen. Die größten Probleme

bereitete die Frage nach dem Zugang zu einem Rechner sowie, im Zusammenhang damit, die Wahl eines geeigneten Programms. Ferner mußte bei der VW-Stiftung ein Antrag auf zusätzliche Mittel zur Finanzierung einer Erprobungsphase gestellt werden.

Aus praktischen Erwägungen fiel schließlich die Entscheidung für das Programmpaket GOLEM von der Firma Siemens. Das Ablocken der Daten konnte zumindest für die Probephase in Dortmund erfolgen, die Verarbeitung wurde vom Siemens-Vertriebsrechenzentrum in Düsseldorf vorgenommen.

GOLEM bestach durch die Möglichkeit der Eingabe auch unformatierter Daten und der nachträglichen bzw. sukzessiven Erstellung eines Thesaurus – beides kam der Struktur der zugrunde liegenden Datenbasis sehr entgegen. Auch das Angebot differenzierter Retrievalverfahren stellte unter der Voraussetzung entsprechender technischer Gegebenheiten (Terminal mit Anschluß an ein Rechenzentrum) ein flexibles Instrument für die Forschung in Aussicht.

Da abzusehen war, daß mit der Intensivverzeichnung der Gesamtbestand in der vorhandenen Zeit nicht würde bewältigt werden können, zumal ja die Erprobung der Arbeit mit GOLEM noch ausstand, mußte ein in sich möglichst sinnvoll abgegrenzter Teilbestand ausgewählt werden. Unter Berücksichtigung der bereits erfolgten Erschließungsarbeiten wurden dazu bestimmt:

- die Korrespondenz des 18. Jahrhunderts vollständig (ab 1750, nur Briefkopien),
- die Kopiebücher des 19. Jahrhunderts bis 1834,
- die Empfängerkorrespondenz der Jahre 1818 – 1834.

Die Begrenzung 1818 – 1834 wurde gewählt, weil dieser Zeitraum die Firmengeschichte von der Geschäftsübernahme von Johann Caspar Harkort V. bis zur Einrichtung der Maschinenfabrik auf Harkorten umfaßt; der größere Teil der Gesamtkorrespondenz ist in diesem Abschnitt enthalten.

Während die manuelle Erschließung der Briefe zügig voranschritt, ergaben sich bei den Probeläufen mit GOLEM verschiedene Probleme. Ein Sample von 200 Briefen aus dem 18. Jahrhundert war als "Spielmaterial" ausgewählt und maschinenlesbar gemacht worden. Bei der Verarbeitung kamen folgende Schwierigkeiten zutage:

- der Zeitaufwand für die Übertragung der Daten in GOLEM-Format erwies sich wegen des etwas umständlichen Steuerzeichensystems als ziemlich hoch;
- durch ungeübte Datatypistinnen kam es zu hohen Fehlerquoten bei der Datenerfassung;
- die Kooperation mit dem Düsseldorfer Rechenzentrum war durch Verständigungsschwierigkeiten aufgrund unterschiedlicher Erwartungshaltungen durch gewisse "Reibungsverluste" gekennzeichnet;
- GOLEM war nicht in der Lage, aus unseren Daten ein annähernd druckreifes Register-Manuskript zu erstellen.

Der Gedanke, die gespeicherten Daten für ein Benutzer-Retrieval verfügbar zu machen, mußte aus finanziellen Gründen vorerst fallengelassen werden; daraufhin wurde die Erprobung des EDV-Einsatzes im Oktober 1980 fallengelassen, der Antrag auf Zusatzmittel bei der VW-Stiftung zurückgezogen. Zu diesem Zeitpunkt waren insgesamt etwa 20.000 Briefe einzeln verzeichnet: die Briefkopiebücher von 1750 bis 1721 vollständig, die Empfängerkorrespondenz von 1818 bis 1825.

Ergebnisse der bisher durchgeführten Erschließung

Trotz des Abbruchs der EDV-Erprobung kann man nicht von einem Mißerfolg des Projekts sprechen: das Grundkonzept erscheint nach wie vor sinnvoll, und die inzwischen im Umgang mit Computern gewonnenen Erfahrungen sollen im Rahmen eines neuen – dann allerdings noch stärker forschungsorientierten – Fortsetzungsprojekts weiter verwertet werden.

Die Kartei der indizierten Briefe, alphabetisch nach Korrespondentennamen geordnet, ist indes archivintern als Findmittel bei Recherchen zu benutzen und bildet den Grundstock für weitere Arbeiten.

Auch die im ursprünglichen Antrag vorgesehene Publikation soll realisiert werden; allerdings wurde dazu auf die alte Konzeption zurückgegriffen: eine Veröffentlichung, die in Registerform erfolgen soll, ist nur sinnvoll, wenn sie sich auf ein einigermaßen zusammenhängendes Textcorpus bezieht, ein Kriterium, dem das Corpus der erschlossenen 20.000 Briefe nur mit Einschränkungen genügt. So ist nun vorgesehen, ein Personen- und Ortsregister aller Hauptbücher, Briefkopiebücher und Memorialien aus den Jahren 1712 bis etwa 1800, also praktisch für das 18. Jahrhundert, zu publizieren. Die Erstellung dieses Registers ist im Rohbau bereits abgeschlossen.

Dem Vortrag von Dr. J. Marschalck folgten einige Fragen, denen sich unmittelbar der Vortrag v a n d e V o o r t s anschloß.

S c h o t e l erkundigt sich, wieviel Personen an dem Verzeichnis der 20.000 Briefe gearbeitet haben.

M a r s c h a l c k gibt sich als einzige Person an.

S c h i m m e l p e n n i n c k fragt, wieviele Briefe noch zu bearbeiten seien.

M a r s c h a l c k antwortet, daß bei einer Gesamtschätzung von 70.000 Briefen also noch 50.000 bearbeitet werden müßten.

V a n d e V o o r t fragt, ob Dr. Marschalck Untersuchungen über die Brauchbarkeit des Golemprogramms angestellt habe.

M a r s c h a l c k berichtet daraufhin, sie habe an einem Kursus bei Siemens in München teilgenommen. Dort werde aber mit einem moderneren Computer gearbeitet, wie er in Dortmund nicht zur Verfügung gestanden habe. Das sei vor allem beim Sortieren der verschiedenen Gruppen von Namen spürbar geworden. Die Korrespondentennamen mußten nämlich gemischt oder eingearbeitet werden mit den indirekt gewählten Namen aus der Korrespondenz. Es gelang dem Computer nicht, die aus zwei Ebenen zugeführten Daten mit verschiedenem Status in einem Gesamtindex zusammenzufassen. Außerdem war es nicht möglich, dieses Programm – ein Retrieval-Programm – dem Benutzer zugänglich zu machen. Deswegen sei das Programm nicht fortgesetzt worden.

DAS MARDOC-PROJEKT: EDV IN DEN NIEDERLÄNDISCHEN MARITIMEN MUSEEN

von Drs. J.P. van de Voort

Entstehungsgeschichte von MARDOC

Bibliotheken, Archive und Museen sind das Gedächtnis der Gesellschaft. Dort sind ungeheure Mengen von Angaben über unsere Vergangenheit aufbewahrt. Die Zugänglichkeit ist ein fortdauerndes Problem, in vielen Ländern auch ein Objekt der Staatsaufsicht. Dies resultiert dann in einer Gesetzgebung und, damit verbunden, einer gewissen finanziellen Garantie, die gewährleistet, daß die gesetzlich geforderten Aktivitäten auch tatsächlich ausgeführt werden können. Diese Regelung gilt auch für die Niederlande; Archive, Bibliotheken und Museen sind hier jedoch nicht gleichberechtigt. Die Archive können bereits seit 1918 gewisse Garantien aus einem Archivgesetz geltend machen. Dasselbe gilt seit 1975 für die öffentlichen Bibliotheken. Über ein Museengesetz wird zur Zeit lediglich nachgedacht. Mit der Zugänglichkeit von musealen Kollektionen ist es in den Niederlanden im allgemeinen recht betrüblich bestellt.

Dies beunruhigte auch eine Anzahl von Direktoren von vor allem maritimen Museen. 1969 faßten sie den Plan, einmal zu untersuchen, ob nicht bessere Aufschlußmethoden für museale Kollektionen zu finden wären als die traditionellen. Sie dachten dabei vor allem auch an den Computer als Hilfsmittel. Ihr Vorschlag fand Gehör beim Ministerium für Kultur, Erholung und gesellschaftliche Arbeit, das eine bescheidene Subvention gewährte, um einen Mitarbeiter anzustellen. Man begann mit der Erstellung eines Literatursystems für die Geschichte der niederländischen Seefischerei, genannt VISDOC: ein Durchsicht-Lochkartensystem mit einer Kapazität von 14.000 Literaturdokumenten. Die Stichworte (Deskriptoren), die sowohl für das Indexieren der Literaturdokumente als auch für das Wiederauffinden benötigt wurden, wurden in einem Verzeichnis festgelegt. Das VISDOC-System wurde 1973 handhabbar. Es faßt nun mehr als 3.000 Literaturdokumente.

1975 wurde beschlossen, das Arbeitsgebiet auf die gesamte niederländische maritime Kultur auszuweiten. Man wollte ein automatisiertes Dokumentationssystem schaffen, das die Angaben über maritime Gegenstände, Abbildungen, Archivalien und Literatur, vor allem in Museen, zugänglich machen sollte. Angesichts des Umfangs des Materials wurde beschlossen, sich vorläufig zu beschränken auf Gegenstände und Abbildungen – Kategorien von Dokumenten, die für Museen am wichtigsten sind. Das maritime Dokumentationsprojekt sowie die Zusammenarbeit mit maritimen Museen, die dieses System realisieren wollen, wird seitdem mit "MARDOC" bezeichnet. Die folgenden Museen arbeiten daran mit: das Rijksmuseum "Nederlands Scheepvaartmuseum" in Amsterdam, das Rijksmuseum "Zuiderzeemuseum" in Enkhuizen, das Maritiem Museum "Prins Hendrik"

in Rotterdam, das Marinemuseum in Den Helder und das Fischereimuseum in Vlaardingen. Daneben sind noch 21 Museen mit maritimen Kollektionen prinzipiell bereit, ihre Kollektionen ebenfalls noch dem MARDOC-System zugänglich zu machen.

Off-line Computerverarbeitung

Das Endziel von MARDOC ist eine Datenbasis, die unmittelbar zugänglich ist (on-line) durch Terminals in den angeschlossenen Museen, sowohl für Eingabe, Befragung und Ausgabe von Informationen. Ein Computersystem ist nur dann sinnvoll, wenn es um sehr große Mengen von Daten geht. Eine Untersuchung hat ergeben, daß lediglich 8 % der maritimen Kollektionen in den niederländischen Museen ausreichend beschrieben ist. "Ausreichend" will hier sagen, daß die Objektbeschreibungen mehr oder weniger die Angaben enthalten, die in der musealen Praxis nötig und sinnvoll sind (Beispiel: Beschreibung eines Schiffsmodells). Dies bedeutet, daß für den größten Teil der maritimen Kollektionen die Angaben noch gesammelt und festgelegt werden müssen. Die Museen verfügen jedoch über zu wenig fachkundiges Personal, um auf Dauer eine größere Produktion von Objektbeschreibungen zu liefern.

Die Vorbereitung eines on-line Computersystems setzt voraus, daß man für eine Reihe von Jahren mit ausreichenden finanziellen Mitteln rechnen kann für die Anschaffung oder Miete von Apparaturen, Programmiermaterial oder die Beschäftigung von Fachleuten auf dem Gebiet der Automatisierung und dokumentarischen Information. Die Museen in den Niederlanden verfügen weder über die Fachleute noch über die notwendigen finanziellen Mittel. Aus all diesen Gründen hat MARDOC beschlossen, vorläufig die volle Aufmerksamkeit zu richten auf das größte Problem, d.h. das Beschreiben der maritimen Kollektionen in den Museen. Um in der Zwischenzeit dennoch über Resultate verfügen zu können – das ist auch psychologisch sehr wichtig – werden die einmal beschriebenen Kollektionen über off-line Computerverarbeitung zugänglich gemacht. Dieses beinhaltet, daß die Objektbeschreibungen periodisch durch ein Computerzentrum zu gedruckten Katalogen und Listen verarbeitet werden. Bei dieser Art von Computerverarbeitung brauchen die Museen selbst weder Automatisierungsfachleute noch Apparate und Programmierbedarf. Off-line Computerverarbeitung ist verhältnismäßig preiswert, wenn es um Drucksachen wie Kataloge und Listen geht. Die Angaben abzüglich der maritimen Kollektionen werden außerdem elektronisch festgelegt, was den späteren Übergang zu einem on-line System erleichtert.

Auf Bitten der Organisatoren dieses Symposiums werde ich nun einige Punkte erläutern, die Sie mir vorgelegt haben:

Vorbereitung des Materials

Welche Vorbereitungen des Materials sind notwendig, bevor die Angaben automatisch verarbeitet werden können?

Bei jedem Dokumentationssystem, manuell oder automatisiert, geht es darum, Dokumente nach bestimmten Merkmalen sortieren zu können. Welche Merkmale für die Selektion in Frage kommen, ist abhängig vom Fachgebiet und von den Forderungen (Fragekatalog) der Benutzer. Es muß nun eine Verbindung hergestellt werden zwischen diesen Merkmalen und den Dokumenten, in denen sie vorkommen. Bei MARDOC werden die Merkmale auf einem Beschreibungsformblatt als Stichworte festgelegt. Dieses Formblatt hat eine festgelegte Nummer: die Inventarnummer des beschriebenen Gegenstandes. Das Formblatt wird in das Computersystem durch Lochkarte oder durch einen Terminal eingeführt. Der Computer produziert anschließend automatisch den Index (bzw. Indizes), d.h. eine Liste von Stichwörtern in alphabetischer Reihenfolge und wirft bei jedem Stichwort die Nummern (= Inventarnummern) der Formblätter aus, in denen die Stichworte vorkommen. Daneben schreibt der Computer den Katalog, d.h. den Text der Beschreibungsformulare in der Reihenfolge der Inventarnummern.

MARDOC verwendet die "record cards" der Museum Documentation Association in England als Beschreibungsformular. Es gibt Formblätter für verschiedene Arten von Gegenständen, für den Inhalt von Abbildungen und für topographische Beschreibungen. Eine andere wichtige Vorbereitung ist die Standardisierung der Terminologie. Die Benutzer der Formblätter müssen dieselben Gegenstände mit denselben Stichworten bezeichnen. Man wird also Listen mit Stichworten (Descriptor) erstellen müssen, mit Verweisung auf Synonyme, Homonyme, verwandte Bezeichnungen usw. Unsere Erfahrung ist, daß dies eine zeitraubende Arbeit ist, die jedoch gleichwohl notwendig ist. Bei MARDOC wurden hierfür Arbeitskreise von Konservatoren gebildet. Einer dieser Arbeitskreise hat drei Jahre an einer Anleitung zum Beschreiben von Schiffsmodellen gearbeitet. Dieses Instruktionbuch umfaßt Stichwortlisten (u.a. Namen von Schiffstypen, Bauarten von Schiffsmodellen usw.), aus denen die Benutzer der Formulare auswählen müssen. Das Handbuch umfaßt auch Instruktionen, welche Angaben in welchen Rubriken des Formulars notiert werden müssen. Man arbeitet jetzt an einem Instruktionbuch zum Beschreiben des Inhalts von Abbildungen (Gemälde, Zeichnungen, Stiche, Fotos).

Art und Umfang des bearbeiteten Materials

Eine Untersuchung nach Art und Umfang der maritimen Kollektionen in 31 niederländischen Museen im Jahre 1978 brachte zutage, daß mindestens 200.000 maritime Gegenstände und Abbildungen in diesen Museen magaziniert sind. Mehr als die Hälfte hiervon sind Abbildungen (Gemälde, Zeichnungen, Stiche, Fotos). Wie erwähnt, sind mehr als 90 % dieser Kollektionen unzureichend beschrieben. Angesichts der sehr beschränkten finanziellen Möglichkeiten von MARDOC können wir nichts anderes tun, als zu versuchen, diese Kollektionen Stück für Stück zu bearbeiten.

Begonnen haben wir mit einer für maritime Museen typischen Kategorie von Objekten, den Schiffsmodellen. Davon sind etwa 3.100 vorhanden. Nachdem die vorerwähnte Anleitung zur Beschreibung von Schiffsmodellen fertiggestellt war, hat ein Marinehistoriker, der gerade sein Studium beendet hatte, 635 Schiffsmodelle in 19 verschiedenen Museen beschrieben. Hierfür wurde ein Formblatt der Museum Documentation Association aus England benützt, das zur Beschreibung von historischen Gegenständen entwickelt wurde. Diese Formblätter sind durch den Computer der MDA zu einem Katalog und zu Indizes verarbeitet worden.

Im kommenden Halbjahr erwarten wir noch ca. 600 Schiffsmodellbeschreibungen und ca. 800 Beschreibungen von maritimen Abbildungen, die durch den Computer verarbeitet werden können.

Apparatur, Computerprogramm

Wie gesagt, verfügt MARDOC selbst nicht über diese Apparatur. Für die off-line Computerverarbeitung arbeiten wir zusammen mit der MDA in England. Die MDA ist eine von den britischen Museen gegründete, nicht kommerzielle Dienstleistungsorganisation, welcher zur Zeit ca. 300 britische Museen angeschlossen sind. Die MDA entwickelt Beschreibungsformblätter, Programme (vor allem computerpackage GOS), gibt Ratschläge und Instruktionen und hilft Museen beim Erstellen ihrer Dokumentationssysteme. Auch die Forschung auf dem Gebiet der Dokumentation und Automatisierung gehört mit zu dem Aufgabengebiet. Mitglieder und Nichtmitglieder können gegen Entrichtung der festgelegten Gebühren von den Möglichkeiten der MDA Gebrauch machen. Ich komme darauf noch zurück.

Unsere Erfahrungen mit der MDA sind bis jetzt positiv. Das Verfahren hat sich wie folgt eingespielt: MARDOC sorgt selbst dafür, daß die Kollektionen auf Formblättern beschrieben werden, und zwar möglichst auf MDA-Formblättern, wengleich dies nicht Bedingung ist. Die Formblätter werden periodisch mit einer Spezifikation bezüglich der gewünschten Kataloge und Listen zur MDA gesandt. Die Computerverarbeitung durch die MDA umfaßt folgende Verfahrensabläufe und Programme:

1. "Data preparation"

Eine Datentypistin macht die Beschreibungen maschinenlesbar mit Hilfe einer Lochstreifenmaschine. Dabei druckt sie vor jeder Information bzw. vor jedem Informationsblock ein sogenanntes "tag" (Identitätskennzeichen) und/oder den sogenannten "Separator" (Trennungszeichen). Der Lochstreifen wird in den Computer "eingelassen" und die Angaben werden auf Magnetband oder -scheibe gespeichert. Die Lochstreifenmaschine liefert gleichzeitig eine Papierrolle, auf der alles ausgedruckt ist, was die Locherin eingegeben hat. Auf dieser Rolle markiert sie die durch sie gemachten Fehler und fügt eventuell Bemerkungen ein.

2. "Source listing"

Der Computer liefert eine "source list", eine Liste mit allen Informationen, wie sie durch die Locherin eingegeben wurden, wobei die vorgenannten Berichtigungen der Locherin bereits berücksichtigt sind. Jede Zeile der source list ist numeriert. Diese Liste erhält der Benutzer nach jedem Durchlauf und kann darauf Fehler markieren.

3. Mit dem BUILD-programma wird der Basistext (s. source list) umgesetzt in eine Standardstruktur: das Standardformat der MDA, das insgesamt ca. 130 Felder umfaßt. Das Standardformat kommt zustande, weil der Computer die obengenannte tags und separators interpretieren kann. Auf der Grundlage dieses Standardformats werden die meisten anderen Bearbeitungen ausgeführt.

4. Das DISPLAY-programm ist bestimmt für die Aufmachung des records (Beschreibung). Es kann die Reihenfolge der Angaben innerhalb eines records, nicht jedoch die record-Reihenfolge in der Liste ändern.

5. Mit Hilfe des KEY- und SORT-Programms werden die records in einer Liste (z.B. Katalog oder Index) in der gewünschten Reihenfolge aufgelistet.

6. Mit dem PAG-Programm wird das lay-out pro Seite erstellt (Paginierung, Textbreite, Köpfe usw.).

7. Der Benutzer behält die source-list und einen Probekatalog und trägt darin die notwendigen Korrekturen bzw. Veränderungen ein. Diese Listen werden danach der MDA zurückgesandt, wo die Korrekturen und Veränderungen berücksichtigt werden. Der Benutzer erhält wieder eine neue source-list sowie die erbetenen Kataloge bzw. Indizes. Falls gewünscht, werden diese Verfahren für neue Korrekturen und/oder Ergänzungen wiederholt.

Das GOS-Programm ist vor allem bestimmt für die Erstellung von Listen, die Beschreibungen mit übereinstimmender interner Struktur enthalten, wie Kataloge, Inventarverzeichnisse, Register, Indizes, Wörterbücher usw. Das Programm ist in der höheren Programmiersprache BCPL geschrieben. Für eine detaillierte techni-

sche Beschreibung des GOS-Programms muß ich Sie auf die Publikationen der MDA verweisen.

Die MDA kann die Kataloge und Listen als Computerdrucksachen, gewöhnliche Drucksachen, auf Mikrofilm oder auf magnetischem Band liefern.

Kosten und Personal

Hinsichtlich der Kosten und des Personals müssen wir folgendes unterscheiden:

- Vorbereitung des Materials, d.h. Angaben über Gegenstände und Abbildungen auf Beschreibungsformblättern eintragen
- Computerverarbeitung dieser Formblätter.

Vorbereitung des Materials

Zunächst ist es notwendig, die Beschreibungsformblätter zu beschaffen. Die MDA-Formulare kosten £ 4,50 je 100 Stück. Weiterhin fallen Personalkosten für das Ausfüllen der Formulare an. Es ist schwierig, hierfür konkrete Beträge zu nennen. Die Zeit, die für das Ausfüllen eines Formblattes benötigt wird, ist von vielen Faktoren abhängig, wie Fachkunde des Ausfüllenden, Art des Objektes, Qualität der bereits vorhandenen Dokumentation, Anforderungen an die Beschreibung. Bei dem bereits erwähnten Test mit der Beschreibung von Schiffsmodellen, erforderte das Ausfüllen des Formulars einschließlich der maschinenschriftlichen Übertragung 62 Minuten. Der Beschreiber war allerdings nicht vertraut mit der in dem betreffenden Museum vorhandenen Dokumentation. Beschreibung durch eigenes Museumspersonal wird weniger Zeit erfordern, nach unserer Erfahrung 15 bis 45 Minuten, je nach Art des Objektes.

Zur Qualität des Personals:

Im allgemeinen wird das Formblatt von oder unter unmittelbarer Aufsicht von Personen ausgefüllt, die ausreichende Kenntnisse auf dem betreffenden Fachgebiet bzw. über die Art der Objekte besitzen, in Museen demnach von Konservatoren. Es ist allerdings möglich, das Ausfüllen von bestimmten Angaben, wie Aufbewahrungsort des Objektes (Depot), Abmessungen der Objekte, Foto- und Negativnummern, administrative Angaben wie Art des Erwerbs, Rechnungsnummer usw. zu delegieren, z.B. auf Verwaltungspersonal.

Kosten der Computerverarbeitung

Aufgrund der Tarife der MDA von September 1981 können wir global die Kosten angeben, die zur Erstellung eines Katalogs (mit neuen Indizes) entstehen. Sie basieren auf einer Sammlung von 1000 Beschreibungsformblättern von i.M. 1000 Zeichen; es handelt sich dabei um ausführliche Beschreibungen.

Die Datentypistinnen der MDA können i.M. 10 Beschreibungen pro Stunde maschinenlesbar machen ("data preparation"). Hierfür werden demnach $100 \times \text{£ } 9,50 = \text{£ } 950$ berechnet. Als Faustregel gilt, daß die Computerverarbeitung einschließlich Ausdrucken ca. 50 % der Kosten des Maschinenlesbarmachens umfaßt. Dies sind nochmals $\text{£ } 425$, insgesamt demnach $\text{£ } 1275$. Nach dem heutigen Kurs des englischen Pfundes würde dies einen Preis von hfl. 5737 für 1000 Objekte oder hfl. 5,74 pro Objekt ausmachen.

Ich hoffe, daß ich Ihnen mit diesen Summenangaben eine ungefähre Vorstellung von den Kosten der off-line Computerverarbeitung geben konnte.

Kosten und Nutzen der Computer-Verarbeitung

Auf die Frage, ob die Erschließung von musealen Kollektionen mit Hilfe des Computers kostengünstiger ist als die traditionellen Methoden, kann ich Ihnen keine Antwort geben. Tatsache ist, daß es den Museen immer noch nicht gelungen ist, ihre umfangreichen und noch immer anwachsenden Kollektionen zugänglich zu machen.

Man muß sich daher wohl nach anderen als den traditionellen Mitteln umsehen. Hinzu kommt, daß heute ein viel größeres Publikum als früher in den Museen Informationen sucht, und dies außerdem durch die Regierung gefördert wird. Wenn die Museen ihre Kollektionen wirklich erschließen wollen, haben sie keine andere Wahl als die Automatisierung. Es sind allerdings große Anstrengungen erforderlich, um das Material so zu bearbeiten, daß eine Automatisierung möglich ist. Der Lohn dieser Mühen liegt freilich darin, daß die Zugänglichkeit durch Kataloge und Indizes sozusagen automatisch erreicht wird.

Schlußfolgerungen

Die maritimen Museen, die in dem Projekt MARDOC zusammenarbeiten, haben sich dafür entschieden, ihre Kollektionen über off-line Computerverarbeitung zugänglich zu machen und im übrigen bessere Zeiten abzuwarten. Eine effizientere und großzügigere Lösung ist für sie wegen der ungenügenden finanziellen Ausstattung, des Personalmangels und unzureichender Automatisierungsfachkunde in Museumskreisen vorläufig nicht absehbar.

Es besteht eine starke Analogie zwischen dem Zugänglichmachen von Archivkollektionen und Museumskollektionen. Auch in Archivkreisen sind Inventare, Kataloge und Indizes die geeigneten Erschließungsmittel. Standardisierung der Terminologie wird darum auch bei Ihnen eine große Rolle spielen. Und ebenso wie die Museen werden auch die Archive fortlaufend durch personelle und finanzielle Probleme betroffen. Ich hoffe daher, daß die Information über MARDOC für Sie von Nutzen sein wird.

Literaturhinweise

Chenhall, R.G.: Museum Cataloging in the Computer Age, Nashville (Tennessee), 1975.

MARDOC-handleiding voor de beschrijving van scheepsmodellen, Vlaardingen, 1977 (= MARDOC-publicatie no. 1).

Porter, M.F.: GOS Reference Manual. Duxford, Museum Documentation Association, 1981 (computer package GOS).

Roberts, D.A., Light, R.B.: Progress in documentation. Museum documentation. In: Journal of Documentation 36 (1980), no. 1, p. 42 – 84 (mit Bibliographie).

Roberts, D.A., Light, R.B., Stewart, J.D.: The Museum Documentation Association. In: Museums Journal 80 (1980), no. 2, p. 81 – 85.

Voort, J.P. van de: Van model naar werkelijkheid. MARDOC voortgangsrapport, Vlaardingen, 1980 (= MARDOC-publicatie no. 3).

Voort, J.P. van de: Maritieme musea hebben wind tegen bij uitvoering MARDOC-project. In: Museumvisie 5 (1981), no. 3, p. 69 – 71.

EDV-ARBEITEN AN DER ARCHIVSCHULE MARBURG UND DEM HESSISCHEN STAATSARCHIV MARBURG IN ZUSAMMENARBEIT MIT DEM RECHEN- ZENTRUM DER UNIVERSITÄT MARBURG

von Inge Auerbach

Staatsarchiv und Archivschule Marburg haben inzwischen eine etwa zehnjährige Erfahrung mit der Anwendung der EDV zu archivischen Zwecken. Die beiden Institutionen lassen sich hier nur als Einheit behandeln, da der Unterricht bzw. die Datenerfassung an der Archivschule durch einen Bediensteten des Staatsarchivs geleitet werden, dieser dann die Weiterverarbeitung übernimmt, die Archivschule jedoch bis 1978 die Kosten der Publikation der Computerausdrucke getragen hat. Lassen Sie uns zunächst auf die Vorgeschichte eingehen, da die historische Entwicklung noch heute für das gewählte Verfahren bei der Anwendung der EDV von Bedeutung ist.

Auf Wunsch des 11. wissenschaftlichen Lehrganges wurde 1971 an der Archivschule Unterricht in Datenverarbeitung eingeführt. Ein Kurs in Fortran IV wurde dem Leiter des Rechenzentrums der Universität Marburg Otto Fröhlich übertragen. Im Laufe dieses Kurses entwickelte Otto Fröhlich auch das Programmpaket HETRINA, mit dem auf dem Rechner der Universität das erste praktische Projekt der Archivschule gerechnet wurde. Die Daten für dieses Projekt wurden im Rahmen einer 14tägigen Ordnungsarbeit der Archivschüler erfaßt. Der Unterricht in Fortran überstieg jedoch schon von Anfang an die Fähigkeiten eines Teils der Archivreferendare bzw. wurde – es handelte sich nicht um ein Prüfungs- oder Wahlprüfungsfach – nicht durch intensives Nacharbeiten des gebotenen Stoffes fruchtbar. Auf Anregung des 16. Referendarlehrganges wurde dieser Unterricht in einer Programmiersprache wieder aufgegeben. Was verblieb, ist die Zusammenarbeit der Archivschule bzw. des Staatsarchives bei allen technischen Fragen mit dem Rechenzentrum der Universität Marburg.

Ab 1972 wurde auch dem gehobenen Dienst, den Inspektoranwärtern an der Archivschule eine Einführung in archivische Probleme der Datenverarbeitung zuteil. Sie wie die Referendare absolvierten eine praktische Übung in archivischer Anwendung der EDV. Aus diesen praktischen Übungen sind die Referendare mit dem 17. Lehrgang entlassen worden, so daß die jüngsten EDV-Projekte allein mit den Inspektoranwärtern durchgeführt werden.

Die praktische Übung bestand für beide Kurse ursprünglich in der Erfassung von Daten nach einem bestimmten Schema aus historischen Quellen (d.h. nicht maschinenschriftlichen), im Ablochen dieser Daten und in einer Beteiligung an den Korrekturarbeiten an den ersten Computerausdrucken. Dieses Programm wurde im Laufe der Zeit reduziert. Heute wird den Archivschülern zwar die Möglichkeit geboten, sich über das Funktionieren des Lochers zu informieren und auch selbst Loch-

karten herzustellen, für die Locharbeiten wurde jedoch eine Hilfskraft gewonnen. Sie sind nur noch bei der Datenerfassung beteiligt.

Aus dieser Vorgeschichte erklärt sich das gewählte Verfahren. Da das Erlernen einer gewissen technischen Fertigkeit im Locher ursprünglich Unterrichtsziel war, entschloß man sich, die Daten auf Lochkarten zu erfassen. Da im Unterrichtsbetrieb mit dem Herunterfallen der Kartenstapel zu rechnen war, empfahl es sich, Fortsetzungskarten zu vermeiden. Ein Datensatz mußte daher mit den 80 Positionen einer Lochkarte auskommen. Der verwandte IBM-Locher 29 ist von der Archivschule gekauft, daher ist heute ein Wechsel auf andere Erfassungsmittel nicht mehr möglich.

Die Rechenarbeiten erfolgten auf einem Telefunkerrechner TR 4 der Universität Marburg. Das Fortran-Programmpaket HETRINA für die Erschließung von personenbezogenen Daten aus den Hessen-Kasselschen Akten zum amerikanischen Unabhängigkeitskrieg war auf diesen Rechner zugeschnitten und wurde für Band 1 – 5 der HETRINA-Serie verwandt. Da die Universität noch während der Erfassung von Daten zu Band 6 dieser Serie einen neuen Computer installierte, ergab sich eine Modifikation des Programms. Ein Teil der notwendigen Operationen gehört nun zu den Standardprogrammen des neuen Telefunkerrechners. Die Anpassung an die neuen Gegebenheiten hat Otto Fröhlich kostenlos in eigener Regie vorgenommen. Da sich diese Arbeiten nur für den einen, noch ausstehenden HETRINA-Band nicht lohnten, wurde das neue, an der Archivschule begonnene Datenverarbeitungsprojekt HESAUS (Hessische Auswanderer) 1840 – 1850 so eingerichtet, daß es mit dem gleichen Programm gerechnet werden kann.

Die Zusammenarbeit mit dem Rechenzentrum der Universität am Ort hat sich bewährt. Die Arbeiten werden dort für uns zu einem verbilligten Sondertarif vorgenommen, so daß die Kosten für die Rechenzeit gemessen an den Gesamtkosten der EDV-Projekte kaum ins Gewicht fallen.

Die Personalkosten sind jedoch im Gegensatz zu den Rechenkosten sehr hoch. Der stundenweise am Locher arbeitende Student (1 – 2 Tage pro Woche) kostet im Jahr etwa 8.500 DM, dazu kommt für die Korrekturarbeiten 3/4 eines Oberratsgehaltes. Es hat sich in den zehn Jahren herausgestellt, daß sich anders als bei neuzeitlichen Akten die Routinearbeiten nicht auf angelernte Hilfskräfte abwälzen lassen. Dies zeigt sich eindeutig an den zahlreichen Lesefehlern, die die Archivschüler noch machen, nachdem sie bereits mindestens

ein halbes Jahr Unterricht in neuerer deutscher Paläographie hinter sich haben, meist auch schon in ihrer praktischen Ausbildung mit handgeschriebenen Quellen in Berührung gekommen sind. Als Folgerung ergibt sich, daß für die Indizierung der Quellen angelernte Kräfte oder jüngere Inspektoren ohne sehr gute Lesefähigkeit nicht in Frage kommen. EDV bedeutet daher für historische Bestände stets das Abziehen qualifizierter Mitarbeiter von anderen Aufgaben zu Routinearbeiten, die ihnen kaum zuzumuten sind. In unserem Fall rechtfertigt sich der Einsatz der EDV zur Erschließung der Akten im wesentlichen nur mit dem Zweck der Demonstration der Einsatzmöglichkeit des Computers im Archiv zu Unterrichtszwecken. Anders als dies auf den ersten Blick scheinen könnte, bedeutet die Beteiligung der Archivschüler an der Erfassung der Daten nur scheinbar eine Arbeitserleichterung. In der Praxis hält das Korrigieren der Lese- und Erfassungsfehler mehr auf als eine erneute Datenaufnahme.

Wie zu erwarten ergibt sich daraus, wie aus Gründen der Arbeitsökonomie, daß der Einsatz der EDV zur Erschließung von Archivbeständen nur in Ausnahmefällen zu empfehlen ist. Für die Bände 1 – 5 der HETRINA-Serie wurden bei einem Einsatz von 25 – 55 Archivschülern pro Band in fünf Jahren insgesamt nur etwa 2 1/2 Meter Akten ausgewertet, für den 6. Band und einen weiteren Meter Akten ist die Arbeitskraft eines Inspektorenlehrgangs und zwei Jahre Arbeit eines Facharchivars für die Korrekturarbeiten investiert worden. Das neue Projekt, HESAUS 1840 – 1850, erschließt in ca. ein- bis anderthalbjähriger Arbeit bei Vereinfachung des Schemas der Datenaufnahme und geringerer Dichte der Daten pro Seite ca. 8 – 10 Meter Akten. Die Datenerfassung erfolgte mit drei Archivschulkursen.

Für die Wahl des Projektes waren zwei Gesichtspunkte ausschlaggebend:

1. eine Erleichterung der Auskunftstätigkeit des Staatsarchivs bei sich häufig wiederholenden Anfragen über bestimmte Bestände,
2. die Möglichkeit einer Publikation der Ergebnisse der EDV-Arbeit durch Auswertung von Quellen, die nicht dem Datenschutz unterliegen.

Gewählt wurden Bestände zu den hessisch- bzw. waldeckisch-amerikanischen Beziehungen, d.h. personenbezogene Akten aus dem Amerikanischen Unabhängigkeitskrieg (HETRINA) und Auswandererakten (HESAUS). Die für den Zeitraum vom 1.1.1840 – 31.12.1850 begonnene HESAUS-Serie erfaßt jedoch nicht nur die Auswanderer nach Amerika, sondern alle, auch die innerdeutschen Auswanderer aus Hessen-Kassel. Die HESAUS-Serie soll für die Jahre 1841 – 1866 fortgesetzt werden.

Positiv wirken sich die Computerprojekte vor allem auf den folgenden Sektoren aus:

1. für die Konservierung der bei genealogischen Anfragen häufig durchzusehenden Akten,

2. für die größere Korrektheit der Auskunftserteilung.

(Erst durch die Erschließung der Akten mit Datenverarbeitung stellte sich heraus, daß die Aktenführung der hessischen Zentralbehörden für die Auswanderungssachen im 19. Jahrhundert mangelhaft war. Ein Teil der Auswanderer nach Amerika findet sich in Aktenbänden unter dem Titel "Auswanderung allgemein", ein anderer unter dem Titel "Auswanderung nach Amerika".)

3. Für die Rekonstruktion der vollständigen Daten über die Auswanderung aus Hessen. Durch das Zusammentragen der Angaben aus verschiedenen Provenienzen sollen die durch Kassation entstandenen Lücken in der Überlieferung wieder gefüllt werden. Im Falle des HETRINA-Projektes lassen sich diese Lücken nicht füllen, doch zumindest die erhaltenen Nachrichten an einer Stelle zusammenbringen.

4. für Antworten auf neue Fragestellungen (Beispiel: für die Lokalgeschichte der Nachweis aller Teilnehmer am amerikanischen Unabhängigkeitskrieg aus einem bestimmten Ort).

Unbrauchbar sind die Computerausdrucke allerdings zur Berichtigung der überlieferten Statistiken über die Zahl der Teilnehmer, der Gefallenen, der Rückkehrer usw. im amerikanischen Unabhängigkeitskrieg wegen der unvollständigen Aktenüberlieferung für einzelne Regimenter oder der Auswandererstatistiken, da Ehefrauen als nicht wehrpflichtig bei den Auswanderern offenbar häufig weggelassen wurden.

5. für die Öffentlichkeitsarbeit des Staatsarchivs.

Die Computerausdrucke werden mit einer deutsch- und englischsprachigen Einleitung in einer Auflage von 400 – 600 Stück von uns vertrieben und verkaufen sich in wenigen Jahren vollständig.

Publiziert werden nur die Indices nach Personennamen, die Indices aller übrigen Meldungen sind für Interessenten im Lesesaal des Staatsarchives zugänglich.

Die Bänder liegen im Rechenzentrum der Universität Marburg bei dem in beiden Serien für die Programme zuständigen Coautoren Otto Fröhlich.

Die Nutzung der Bänder durch Dritte wirft urheberrechtliche Probleme auf. Auf den Akten des Staatsarchivs Marburg selbst liegen keine Urheberrechte wie durch Erlaß des Hessischen Kultusministers festgestellt wurde. EDV-Programme wie das HETRINA-Programmpaket und seine Weiterentwicklung sind zur Zeit nach deutschem Recht urheberrechtlich nicht geschützt. Urheberrechte liegen jedoch auf den Publikationen. Nach nicht unumstrittener Ansicht des hessischen Kultusministeriums haben Archivare bei Arbeiten, die im Dienst entstehen, keinen Anspruch auf den daraus erwachsenden Gewinn. Als Autoren sind sie jedoch verpflichtet, nach gültiger deutscher Rechtsansicht keine Schritte zu unternehmen, die den Absatz der einem bestimmten Verleger überlassenen Werke schädigen. Den Druck haben bisher die

an dem Publikationsfonds der Archivschule Marburg beteiligten Länder und der Bund finanziert, ab Band 6 und für die HESAUS-Reihe übernimmt die Historische Kommission für Hessen die Publikation. In beiden Fällen kommen etwaige geringfügige Gewinne den Stellen zugute, die Auslagen getragen haben, die für die Erstellung der Arbeiten notwendig waren, der Archivschule bzw. dem Land Hessen. In den Verkaufspreis sind bisher die Kosten für die Rechenzeit, die Materialkosten und

vor allem die Personalkosten nicht einbezogen worden. An einem Raubdruck der sich gut verkaufenden Handbücher können vor allem amerikanische genealogische Gesellschaften interessiert sein. Um einen Schaden, etwa durch Verhinderung von Neuauflagen, von den Herausgebern abzuwenden, wurde amerikanisches Copyright bei der Copyright-Abteilung der Library of Congress, Washington, D.C. beantragt, allerdings ohne Erfolg, da der Termin für die Anmeldung verstrichen war.

Nach dem Vortrag von Dr. I. Auerbach ergab sich die folgende Diskussion.

R o m e y k leitet mit der Bemerkung, daß doch jeder Archivar davon träume, daß seine Findbücher als Raubdruck in die Welt gingen, zur Diskussion über.

R i c h t e r i n g meint, diese Art der Erschließung und Publikation bewältigten nicht nur schneller und besser Anfragen, sondern böten zudem noch den Vorteil, daß die Originalakten konservatorisch geschont würden. Schonung der Originale müsse bei der Entscheidung über die Art und Weise der Erschließung von Archiven in die Erwägung einbezogen werden. Außerdem habe die Erfahrung gezeigt, daß die Erschließung mehr Benutzer anziehe. Es gebe jedoch auch Archive und Bestände, die den Arbeits- und Kostenaufwand, den die Erschließung durch EDV fordere, nicht rechtfertigten.

K e t e l a r stellt die Frage, ob man in Marburg diesen Weg wieder wählen würde, falls sich die Chance dazu nochmals biete.

A u e r b a c h erklärt, daß es sich stets um eine Arbeit handeln müsse, die in einer beschränkten Zeit abzuschließen sei, um den Archivreferendaren sichtbar zu machen, daß ihre Arbeit tatsächlich auch von den Benutzern in Anspruch genommen werde.

K e t e l a r ergänzt, man habe offensichtlich in Marburg – wie übrigens auch in den Niederlanden – versucht, den Archivaren über eine praktische Übung elektronische Datenverarbeitung beizubringen und habe sie doch nur gelehrt, Karten zu lochen.

A u e r b a c h meint, daß die Eingabe von Daten ebenfalls Aufgabe eines Archivars sei. So werde man mit den Vorbereitungsarbeiten und mit bestimmten Schwierigkeiten der EDV-Praxis vertraut. Bei der Eingabe der Daten müsse man Namen und Daten identifizieren sowie Orte ausweisen können. Dies alles seien Arbeiten, die man nicht auf andere Kräfte abschieben könne.

K e t e l a r beharrt dennoch darauf, daß es sich um eine paläographische Übung handle, keineswegs um eine spezifische EDV-Übung.

R o m e y k versucht, die divergierenden Auffassungen in Einklang zu bringen, man müsse nämlich bestimmte paläographische Kenntnisse haben, um Daten richtig erfassen zu können, z.B. das verschnörkelte "g", man vergrößere auch seine landeskundlichen Kenntnisse und erlange schließlich Einsicht in lokal und regional verbreitete Familiennamen. Zwar werde in gewissem Sinne die Arbeitskraft der Referendare ausgenutzt, aber als einstmals Betroffener halte er, Romeyk, dies auch deswegen für vertretbar, weil man in überschaubarer Frist das Arbeitsergebnis zu sehen bekam. Ob dies aber eine angemessene Arbeit für Personen mit Dokortitel sei, sei eine andere Frage.

A u e r b a c h fügt noch hinzu, als man damals in Marburg mit dieser Arbeit begonnen habe, seien nur Indizierungsprogramme durchgeführt worden, die Er-

fahrung sei gering gewesen, und die technische Entwicklung habe im Vergleich zu heute noch am Anfang gestanden. Es sei auch ein Ergebnis dieser Arbeit gewesen zu verdeutlichen, daß die Arbeit mit EDV zunächst eine Mehrarbeit bedeute und nicht immer angenehm sei. Die Schule habe vermitteln wollen, daß man zwar Daten verarbeiten könne, aber auch, daß man sich sehr gut überlegen müsse, ob der Mehraufwand dem angestrebten Nutzen entspreche.

R o m e y k erläutert noch die technischen Beschränkungen: die Rechenkapazität des damals in Marburg zur Verfügung stehenden Computers sei angesichts der heutigen Begriffe sehr gering gewesen.

Nach diesen Fragen zum Referat von Dr. Auerbach erfolgte wieder eine allgemeine Diskussion.

V a n D r i e l fragt sich, ob bei der Erschließung von älteren und großen Beständen die Datenerfassung nicht einen Engpaß darstelle.

F r a u M a r s c h a l c k kommt zurück auf die Diskussion, die dem Vortrag von Frau Auerbach galt, und fügt hinzu, daß die Arbeit gar nicht langweilig zu sein brauche, man habe eben selber die Akten vor Augen und in den Händen und sehe das Fortschreiten der Arbeit. Sie meint auch, daß man keine unqualifizierten Personen an die Erarbeitung setzen sollte. Der Einsatz von EDV lohne sich bei allen Archiven, die einen großen wissenschaftlichen Wert hätten, eben für Bestände, die aktuell seien und eine intensive Nutzung erwarten ließen.

K e t e l a r glaubt, wenn man sich einmal entschieden habe, ein Archiv zu erschließen, bestehe nur die Wahl zwischen der traditionellen oder EDV-Erschließung, und diese falle stets zum Vorteil der EDV aus. Deswegen sei die Fragestellung durch van Driel eigentlich nicht richtig.

Nach Ansicht von R o m e y k wiche man hier etwas vom Thema ab, nicht diese Entscheidung stehe zur Diskussion, sondern die Erschließungsarbeit an sich. Er gab zu bedenken, daß man sich möglichst im voraus darüber klar sein müsse, wie sich die Erschließung entwickle und ob die Kosten vertretbar blieben.

K e t e l a r erwähnt noch einmal das Mardoc-Projekt als ein gutes Beispiel des beschränkten Einsatzes der EDV. Ein Teil des Bestandes sei beschrieben und vorläufig durch ein Repertorium erschlossen. Wenn einmal die Finanzlage besser und die Computer billiger geworden seien, könnten auch die gespeicherten Daten weiter verwandt werden.

J a n s s e n möchte auf die Kostenfrage zurückkommen und führt aus: Die Kosten, die zur Vorbereitung eines Erschließungsprojektes und bei der Aufbereitung der Daten für den Computer anfielen, ergäben sich auch sonst und dürften eigentlich gar nicht zu den spezifischen EDV-Kosten gerechnet werden. Der Kostenvergleich könne erst gemacht werden, wenn die auch sonst üblichen konventionellen Arbeiten abgeschlossen seien, und der

Einsatz des Computers in Vergleich mit der noch zu erledigenden manuellen Verarbeitung gesetzt werde.

R i c h t e r i n g vertritt die Meinung, daß der EDV-Einsatz im Museumsbereich sicher lohnender sei, weil der Wert der Erschließungsmodelle, aber auch die Anfragen danach höher lägen als bei der Masse der Akten.

S c h i m m e l p e n n i n c k erwidert, daß der Unterschied wenigstens prinzipiell nicht so groß sei.

R i c h t e r i n g verweist auf den allein zahlenmäßig großen Unterschied zwischen Archiven mit Akten und Museen mit z.B. Schiffsmodellen.

P e t r y kommt nochmals auf das Kostenproblem zurück. Im Verlaufe der Diskussion sei eine Einsparung der Schreibarbeiten bis zu 25 % genannt worden. In Düsseldorf habe man feststellen können, daß die Einsparungen bei Schreibarbeiten bis 30 % im Vergleich zur manuellen Verarbeitung betrügen, und seines Erachtens könne kein Zweifel daran bestehen, daß die Erschließung mittels EDV rationeller gehe. Aber die Fragen der Benutzer hätten sich in den letzten 10 Jahren qualitativ geändert und quantitativ vervielfacht. Vor allem frage man nach personen- und ortsbezogenen Daten, die normalen Findbücher reichten dafür nicht aus. Ohne die EDV könnten die Archivare in Zukunft immer weniger leisten, da sie zu sehr durch persönliche Einzelberatung beansprucht würden.

R i c h t e r i n g bejaht dies und fügt hinzu, daß der Benutzer weniger Ausbildung und Vorkenntnisse als früher über die Archive hat, aber gleichwohl soviel als möglich aus den betreffenden Archiven herausholen möchte.

P e t r y warnt vor der Gefahr, daß man vor lauter Beratung nicht mehr an seine eigentliche Arbeit komme und davon entfremdet werde. Er verweist auf die archivische Aufgabe, auch das Massenschriftgut zu bewältigen und erwähnt, daß in Düsseldorf jährlich soviel Schriftgut hinzuwachsen wie in 1900 der ganze Archivbestand ausgemacht habe.

S c h i m m e l p e n n i n c k schließt die Diskussion und stellt als besonders interessant die edukativen Aspekte heraus, die bei dem Projekt in Marburg erzielt worden seien. Er spricht die Hoffnung aus, daß auch in den Niederlanden mit einem ähnlichen Projekt begonnen werden möge, vorerst sei dafür aber noch kein Geld vorhanden.

Nach der Sitzung wurden die Gespräche in kleineren Kreisen während des Zusammenseins im Rijksarchief und des Abendessens (auf Einladung des Westfälischen Archivamts) fortgesetzt.

FREITAG, 23. OKTOBER 1981

Die dritte Sitzung wurde moderiert von Dr. O p p e l und Dr. K e t e l a a r. Es berichteten zuerst Dr. s. P. S c h o t e l (Gemeentearchief Dordrecht), H. P. N e u h e u s e r (Archivberatungsstelle Rheinland) und R. R i e b s c h l ä g e r (Rechenzentrum des Landschaftsverbandes Rheinland) über ihre Erfahrungen mit EDV im Archiv.

INCREASING THE ACCESSIBILITY OF THE NOTARIAL RECORDS OF DORDRECHT 1860 – 1869

by P. Schotel

At the municipal record-office of Dordrecht this "notarial project" started in 1977. Its object was to increase the accessibility of the records of the notaries in Dordrecht over the period 1860–1869. This was done in accordance with the rules drafted by a special committee of the Netherlands Archives Council. The various notarial instruments, e.g. testimonial, division of estate, mortgage, inheritance etc. are extracted in accordance with these fixed rules and in this way all the relevant information of the notarial instruments is embodied, i.e. at least the names, the professions, the addresses and the civil status of all the people involved in the action recorded in the instrument. If there are mandataries, the names etc. of their principals are recorded, further the position, names and cadastral number of the real estates and the amount of money involved in the transaction. In case of a public auction of movables, its names, amount and price are given, lists of inventories are not copied but are mentioned. These data are typed on standardized forms, on which every position is numbered and has its particular coded meaning. The number of positions varies with each kind of instrument, but no more than 39 positions are used.

The extraction of the instruments for the period 1860–1869 has been finished now (November 1981), with a total of 8440 instruments. Indices for names, professions, cadastral numbers in and outside Dordrecht and names of houses, sites and ships have been made for 50 of the total of 73 notarial protocols, i.e. of every notary for each year from 1860 up to and including 1869. The time needed for the extraction of a notarial instrument varies between an hour and two days, depending on its complexity and the number of data it contains; especially divisions of estates and inheritances may take a lot of time. The number of employees working on this project changed considerably for various reasons, so an estimate of the time spent on this project cannot be given, but based on the experience of the past six months, the assumption can be made that two experienced employees can extract 10 years of notarial records within 1 1/2 year, not including the manufacturing or the indices.

The object of this project is twofold: 1. a greater accessibility of the notarial records; 2. the protection of the originals, because the researcher is provided with the extracts on the standardized forms and not with the originals.

Hitherto the indices were made by hand and the only way to do this was by working on the notarial protocols one by one, so in the end 73 separate indices would be the result, which is not an ideal situation. Computerizing

these indices, i.e. making one general index, was our original aim. After checking the possibilities of a micro-processor on a commercial basis, we had to apply to a greater system on a non-commercial basis. First because even a micro-processor is too expensive; secondly no micro-processor had a sufficient memory capacity. Therefore we sought contact with the History Department of Erasmus University at Rotterdam in order to attain some kind of cooperation between this department and the Municipal Record-Office of Dordrecht, for in one of his lectures professor H. van Dijk had shown an interest in the possibilities of the use of a computer for historical research. He was interested indeed, not, however, in the computerizing of the indices, but in the electronic handling of all the data embodied on the extract forms of the notarial instruments. After consultations with the Computer Centre of Erasmus University the following arrangements were made, the Record-office of Dordrecht being responsible for the input of the data, and the University of Rotterdam for the hardware and software. It should be noted that we have no experience with the procedure described below, because the microprocessor is not operational as yet. "Dordrecht" borrows a microprocessor from "Rotterdam", type ITT 2020, and the data are typed on it. When the floppy disk is full, a printed output is made, with which the checking of the input is more easily performed. After the correction of the input the floppy disk is sent to the Computer Centre at Rotterdam and transcribed on a greater system with sufficient memory capacity. In this way all the data embodied on the extract forms will be available in the computer memory and electronic data handling be made possible. We are to receive a computer print of all the extract forms and the general index on names etc. for all the notaries of the period 1860–1869. The property of the electronically handled data remains in the hands of Erasmus University. After finishing this project, the continuation of this cooperation has to be judged.

This cooperation enabled us to start a computer project with all its possibilities, which otherwise, for lack of money and know-how, would have remained a utopia for a long time.

So far two conclusions can be drawn: 1. record-office, maybe with the exception of the bigger ones, should not buy the hardware and software at its own expense, because at this moment the cost is too high and in the near future a sharp fall in prices can be expected; 2. whatever the development of the cost of hardware and software may be, the greatest problem will be to find people who are able to perform the task of processing the records necessary for the electronic handling.

VERZEICHNUNG VON KOMMUNALAKTEN MIT HILFE DER ADV

von Hanns Peter Neuheuser

Die anschließenden Darlegungen* sind weniger als methodisch-theoretische Überlegungen anzusehen, sondern sind Ausfluß des gesetzlichen Auftrages der Archivberatungsstelle Rheinland (ABSt) einerseits und Antwort auf die derzeitigen, also aktuellen und konkreten Probleme im nichtstaatlichen Archivbereich andererseits. Zur Vereinfachung haben die Teilnehmer dieses Symposiums zwei Papiere erhalten, die sowohl kurz Aufgaben und Tätigkeiten der ABSt als auch "10 Thesen zur gegenwärtigen archivischen Problemlage der Kommunalarchive im Landesteil Nordrhein" vor Augen führen wollen.

Die Kürze der mir zur Verfügung stehenden Zeit erspart mir, die auf diesem Symposium bereits vorgetragenen allgemeinen Erfahrungen im Umgang mit der Automatisierten Datenverarbeitung (ADV), soweit sie auch meine sind, zu bestätigen. Heute geht es eher darum, einzelne Projekte skizzenhaft vorzustellen.

In den folgenden Ausführungen möchte ich Ihnen die Überlegungen und erarbeiteten Hilfsmittel vorstellen, die Gegenstand der Arbeit der ABSt innerhalb knapper Jahresfrist gewesen sind, und zwar im Hinblick darauf, wie dokumentarische Methoden und technische Möglichkeiten der ADV im Archivbereich, speziell in den Kommunalarchiven, eingesetzt werden können. Bereits 1950 hatte Meiner die Einbeziehung der Archive in das "Gesamtgebäude der Dokumentation" gefordert (NfD 1 [1950], S. 123), ohne daß diese Forderung nach unserer Erkenntnis Konsequenzen für das kommunale Archivwesen erbracht hätte.

Angeichts der besonderen Aufgabendefinition einer Beratungsstelle innerhalb einer konkreten Archivlandschaft mit dringenden, aktuellen Problemen mußten sich die Lösungen als praktikable Hilfsangebote erweisen. Fertig zu erwerbende, jedoch zu umfangreiche und komplizierte Programme verfehlten daher ihre durch vorgebliche methodische Perfektion gekennzeichnete Faszinationswirkung. Die ABSt vermag heute (vorerst) den Kommunalarchivaren ein Instrument zu offerieren, das zunächst nichts anderes wollen kann, als die alltägliche Praxis zu erleichtern.

Durch den substantiellen Vorteil der ADV, nämlich die jederzeitige Verfügbarkeit und die beliebige Verknüpfbarkeit einmal gespeicherter Daten, ergibt sich natürlich eine Erweiterung unseres Vorhabens über die Lösung eines archivischen Einzelproblems hinaus, hin zu einem Informationssystem – COMIS genannt –, dessen Erweiterungsfähigkeit nahezu unbeschränkte Möglichkeiten offenläßt.

COMIS als Gesamtsystem untergliedert sich in weitere Systemteile, die mit ihren Abkürzungen hier kurz erwähnt werden sollen:

1. Gesamtbezeichnung COMIS – Communal Information System

2. Bezeichnung der Systemteile

- COMADES – Computer-Assisted System for the Description of Documents in Communal Archives
- COMPLAN – Klassifikationssystem auf der KGSt-Grundlage
- COMTHES – Kommunal-Thesaurus
- COMARS – Computer-Assisted System for Analysis and Retrieval in Communal Archives (Programm für das information retrieval)
- COMIX – Indizierungsprogramm als Teil von COMADES oder selbständig zur retrospektiven Indizierung fertiger Findbücher
- COMSTAT – Statistisches Programm zur methodischen Analyse von COMIS-Ergebnissen
- COMPLOTT – Graphische Umsetzung analytischer Auswertungen
- COMNET – Verknüpfungsprogramm aller Archive resp. der Dokumentenbeschreibungen und der Analysen

Hierzu ist nachzutragen, daß die Systemteile sich in unterschiedlichen Entwicklungsphasen befinden, vom Planungsstadium bis hin zur vollen Anwendung nach abgeschlossener Testreihe. Dabei ist die Weiterentwicklung der COMIS-Struktur, was zusätzliche Systemteile für weitergehende Verknüpfungen etc. angeht, durchaus möglich.

Unabdingbare Voraussetzung für jedes weitere Operieren, Vergleichen und Zusammenstellen von Daten bildet freilich ihre Erhebung, die Datenermittlung, sprich: das Verzeichnen der Dokumente. Deshalb bildet COMADES sozusagen das Zentrum der COMIS-Struktur.

Je intensiver und extensiver dieser aus arbeitsökonomischen Gründen lediglich einmal durchzuführender Vorgang gestaltet wird, desto mehr Möglichkeiten der Datenverarbeitung eröffnen sich. Einfachheit in der Handhabung und Vielseitigkeit in der Auswertung bilden die obersten Prinzipien, wobei als *conditio sine qua non* vorgegeben war, an der archivwissenschaftlichen Grundmethodik des Verzeichnens nichts zu ändern.

Für die Datenermittlung wird von allen Anwendern einheitlich der gleiche Belegbogen verwendet, in nächster Zukunft zusätzlich und parallel ein dezentrales Datenerfassungsgerät.

* Der Redetext erschien mit den Anlagen als Separatdruck bei der Archivberatungsstelle Rheinland, Köln.

An dieser Stelle möchte ich die einzelnen Bestandteile, d.h. die Gliederung der erhobenen Daten nur jeweils kurz ansprechen, zumal die Deskriptionselemente an sich im Kreise dieses Symposiums keiner Erläuterung bedürfen.

Die Einzeldaten definieren sich durch die Einteilung in Kategorien und Zeilen, welche letztere höchstens 48 Zeichen aufnehmen können. Ansonsten sind zu unterscheiden die Kategorien für die formale Erfassung (d.h. Zahl-Nummer, Signatur, Laufzeit, Fremdprovenienz, historische Signatur und Dokumentenart), die Kategorien für die inhaltliche Erschließung sowie die Kategorien für die Indices. Für die Daten der letztgenannten Kategorien steht der meiste Platz zur Verfügung. Zur Erläuterung des Archivalieninhaltes durch freiformulierten Text können 100 Zeilen benutzt werden, was einem Raum von 4800 Zeichen entspricht. Die – entsprechend der erst kurzen Anwendungszeit natürlich geringe – Erfahrung der ABSt hat gezeigt, daß ein Belegbogen, ein Vordruck, wenn er nach praktikablen Gesichtspunkten aufgebaut ist, den Verzeichnungsvorgang durchaus nicht, wie gelegentlich befürchtet wird, unzulässigerweise einschränkt und behindert, vielmehr eine gründlichere und überlegtere Dokumentenanalyse anregen kann.

Die gleiche extensive Exzerpierung wie für die Freitextdeskription ermöglichen die Kategorien für die Personen- und Ortsindices mit dem Angebot, maximal 100 Personen- und 100 Ortsnamen je Dokument auszuwerfen. Schließlich soll der jeweilige Index nicht nur möglichst viele "aktenkundig" gewordene Namen des Archivzusammenhangs nachweisen, vielmehr durch die Kumulierung mit anderen Indices ein breites Fundament einer überörtlichen Datenbank schaffen.

Ebenfalls 100 Zeilen sind für Bemerkungseintragungen in Kategorie 90 vorgesehen, die zusätzliche Erläuterungen zu Dokumententeilen (Firmenbriefköpfe, Siegelabdrücke, Stempel, Autographen etc.) festhalten. Hierdurch wie schon durch die Erfassung codierter Dokumentenarten wird eine Suchstrategie ermöglicht, die nicht dem Klassifikationsschema unterworfen ist, eine neue Möglichkeit der Recherche speziell für Querschnittsfragen. Im Hauptteil des Belegbogens (Kategorien 50 – 90) finden somit nicht weniger als 19.200 Zeichen Platz. Damit wird ein Informationsfundament bislang nicht bekannter Breite geschaffen.

Die einfache Handhabung des Belegbogens sowie eine Erleichterung der Datenverarbeitung, d.h. die wenig komplizierte Identifikation der Daten, wird erreicht durch die höchstmögliche Kategorisierung der Eintragungen. Die Anzahl der Steuerzeichen wurde zugunsten der annotierten Deskriptoren auf ein Mindestmaß reduziert.

Die Einzelheiten der Kategorien können die Symposiumsteilnehmer der ihnen vorliegenden Benutzeranleitung zu COMADES entnehmen. Die Herausgabe eines COMIS-Handbuchs mit detaillierten Angaben zu allen Systemteilen ist in Lose-Blatt-Form in Vorbereitung.

Heute möchte ich die Aufmerksamkeit nur auf wesentliche Dinge lenken, die mit dem Sortierprozeß zusammenhängen und die die Flexibilität und Erweiterungsmöglichkeiten von COMIS verdeutlichen sollen.

Wichtig ist zunächst die Tatsache, daß bereits beim Verzeichnungsvorgang die beiden ersten Stellen eines vorgegebenen Klassifikationssystems einzutragen sind. Das System ist mit dem allgemeinen Aktenplan für Kommunalarchive – COMPLAN – identisch.

Des weiteren von großer Bedeutung ist Kategorie 40 mit der Angabe eines überschriftartigen Betreffes. An dieser Stelle wird in weiterer Zukunft der Deskriptor eines Thesaurus für den kommunalsprachlichen Wortgebrauch – COMTHES – erscheinen. Die weiteren Bezeichnungen dienen mit Hilfe eines Keyword-in-context-Systems dem Aufbau eines permutierten Sachindex.

Die in den Rechner eingegebenen Daten werden in folgender Reihenfolge sortiert:

1. nach dem Klassifikationszeichen
2. innerhalb dessen alphabetisch nach den Betreffen
3. innerhalb des gleichen Betreffs chronologisch.

Die Zuweisung der Datensätze an ihre endgültige Stelle geschieht – solange der Thesaurus nicht einsatzfähig ist – von Hand im Rahmen des Änderungsdienstes.

Aufgrund der gespeicherten Daten sind folgende Ergebnisse als Standard-Outputs vorgesehen:

1. Der konventionelle Hauptteil eines Findbuchs in der Gliederung der Klassifikation, d.h. nach COMPLAN.
2. Index der Personennamen.
3. Index der Ortsnamen, teilweise durch Dachbegriffe weiter untergliedert, wobei gleiche Einträge nur einmal erscheinen und auf die zugehörigen Signaturen verweisen.
4. Sachindex, durch permutierte Wortprozessierung des "Betreffs".
5. Liste je Dokumentenart zur Auffindung versprengter Archivalien bzw. zur Querschnittsermittlung (z.B. Fotos, Personenlisten, Protokolle etc.).
6. Konkordanz zwischen alten Aktenzeichen und endgültiger Signatur.
7. Aufkleber für Akten mit endgültiger Signatur zur Erleichterung der sonst manuell und nicht selten fehlerhaft durchgeführten Umsignierung.
8. Liste der für die Benutzung gesperrten Archivalien als Hilfe für den Benutzerdienst.
9. Aufkleber für Kartons mit Hinweis auf gesperrte Archivalien als Hilfe für den Magazindienst.
10. Liste der restaurierungsbedürftigen Archivalien als erste Orientierungshilfe für den Werkstattbereich.
11. Liste je Bemerkungsart, z.B. als Hilfe zur Vorbereitung von Ausstellungen oder zur Bebilderung von Publikationen.
12. Liste der Fremdprovenienzen.

13. Inhaltsverzeichnis des Findbuchs mit Verweis auf die einzelnen Abteilungen und die Listen, nach Seitenangaben.

Soweit die vorgesehenen Standard-Ausdrucke, die ohne weitere manuelle Arbeit und ohne zusätzlichen Programmierungsaufwand herzustellen sind. Hierneben kann man sich natürlich eine fast unbegrenzte Zahl von Verknüpfungsmöglichkeiten dieser Listen sowie beliebige Sortierweisen der gespeicherten Daten vorstellen. Verbindet man etwa gleiche Klassifikationsebenen aus verschiedenen Archiven miteinander, so wird man leicht die Fähigkeit des Systems erkennen, nahezu flächendeckend Fragestellungen und Forschungsvorhaben zu beantworten.

Ferner ist COMIS in der Lage, die digital in Listen zusammengestellten Daten in analoge Schaubilder umzusetzen. Ein automatischer, über den Rechner gesteuerter Zeichner, der Plotter, kann mit dem Programm "COMPLOTT" Datenmengen nach unterschiedlichen Gesichtspunkten in Graphiken, Tafeln, Kurven, Säulen etc. veranschaulichen oder Aussagen in die geographische Landkarte unseres Sprengels eintragen.

Durch das statistische Teilsystem COMSTAT bietet sich die Chance, alle Verarbeitungsvorgänge methodisch zu analysieren und Erfahrungswerte zu objektivieren und mit anderen Projekten im Erfahrungsaustausch zu vergleichen.

Soviel zu den Einzelheiten des erarbeiteten Programms. Zusammenfassend lassen sich die Bestrebungen vielleicht in folgenden Postulaten darstellen, in den Anforderungen an die Leistungsfähigkeit des Systems:

1. einfache Handhabung als wirkliches Hilfsinstrument,
2. Gewährleistung gleichmäßiger Verzeichnung in verschiedenen Projekten, d.h. Erhöhung der Qualität der Findbücher,
3. Hilfsmittel zur Bewältigung der Dokumentenmassen,
4. dokumentarische Vielseitigkeit der Daten-Ausschöpfung,
5. Hilfsmittel bei find- und lagertechnischen Schwierigkeiten, bei der Magazinierung, bei der Revision,
6. Benutzerfreundlichkeit bei selbständiger Recherche,
7. Anwendbarkeit bei Querschnittsfragen für Wissenschaft, Forschung und Öffentlichkeit, zumal der archivfremden Benutzer,
8. Kumulierfähigkeit der Systemteile.

Wir sind davon überzeugt, daß das geschilderte System ein solches Hilfsangebot ist, welches dem Kommunalarchivar an die Hand zu geben, Aufgabe der ABSt ist.

Freilich kann das System nicht alle Fragen des archivistischen Alltags lösen, wie wir bereits in unseren Pilotprojekten erfahren konnten. Derzeit befinden sich vier Projekte im fortgeschrittenen Stadium, nämlich die Gemeindearchive Schermbeck, Kranenburg und Much sowie das Stadtarchiv Xanten. Im Rahmen der Erprobung haben wir mit den Akten, die nach dem

Zweiten Weltkrieg angelegt wurden, begonnen, um in der Phase des Kennenlernens einer zum Teil neuen Technologie nicht zusätzlich durch paläographische Schwierigkeiten behindert zu werden; COMIS ist selbstverständlich auch für ältere Bestände konzipiert. Die Bestände unserer Pilotprojekte sind von unterschiedlichem Umfang und reichen wohl von ca. 1500 bis 3000 archivalische Einheiten pro Bestand bei drei bis vier Beständen pro Archiv. Die Datenermittlung, d.h. das Ausfüllen der Belegbögen erfolgt durch drei Mitarbeiter der ABSt, die auch den direkten Kontakt zum Rechenzentrum herstellen, so daß Teilausdrucke eines vorläufigen Findbuches jederzeit gedruckt werden können, um die Funktionsfähigkeit der örtlichen Schriftgutverwaltung zu gewährleisten. Die genannten Mitarbeiter sind Archivare des gehobenen Dienstes und treten an die Stelle der nicht vorhandenen hauptamtlichen Ortsarchivare.

Inzwischen liegen eine Reihe weiterer Anfragen aus rheinischen Archiven vor, welche die Übernahme beantragen. Darunter befinden sich auch hauptamtlich geleitete Archive. Die Einweisung der ersten externen Anwender, d.h. hauptamtlicher Archivleiter ist bereits erfolgt, so daß die ADV-unterstützte Verzeichnung in einer Reihe weiterer rheinischer Kommunalarchive noch in diesem Jahr aufgenommen werden kann.

Die COMIS-Anwender treffen sich regelmäßig in einer Arbeitsgemeinschaft, damit beide Seiten, die Systementwicklung und die Systembenutzung, voneinander und miteinander lernen und speziell für die Anwender eine Identifikation mit einem gemeinsamen Vorhaben heranwächst.

Die Zeit läßt es leider nicht zu, Ihnen weitere Informationen aus der theoretischen und praktischen COMIS-Arbeit zu vermitteln. Details von technischer Seite werden im übrigen im Anschluß an meine Ausführungen von Herrn Riebschläger vom Rechenzentrum des Landschaftsverbandes Rheinland, Köln, mitgeteilt.

Lassen Sie mich nur abschließend und auf Wunsch der Veranstalter dieses Symposiums einige Schlußfolgerungen aus der COMIS-Arbeit versuchen. Wenn es wegen der zugegebenermaßen sehr kurzen und beschränkten Erfahrung überhaupt erlaubt ist, bereits heute zu einem Zwischenergebnis oder gar zu weitergehenden Postulaten zu kommen, so vielleicht, weil man unter Umständen sich noch die Unbefangenheit gegenüber den Phänomenen bewahrt hat. Aus dieser Sicht könnte man aufgrund der praktischen Arbeit folgende sozusagen archivistische Forderungen formulieren:

1. Erkennbar wurde ein fundamentales Sprachproblem; der Archivar muß eine verstärkte terminologische Kontrolle über seinen Verzeichnungstext ausüben. Er selbst muß sich stärker um die Partizipation an der dokumentarischen und ADV-Fachsprache bemühen.
2. Kenntnisse rein archivistischer Methoden genügen nicht mehr. Die Ausbildung des Archivars muß sich mehr als bislang an der komplizierter werdenden Arbeit speziell im Kommunalsektor orientieren.

3. Die Intensität der Verzeichnung darf nicht mehr nur im Blick auf den Einzelbestand erfolgen, sondern muß die mögliche Verknüpfung der Datenbestände berücksichtigen. Insofern ist den Archivaren eine eher dokumentarische Mentalität zu wünschen.
4. Der Archivar muß offener werden für die Bearbeitung der verschiedenen Quellenarten, damit die Daten z.B. einer meist vernachlässigten zeitgeschichtlichen Sammlung nicht „verlorengehen“. Dies gelingt mit verschiedenartigen Erschließungsmethoden.
5. Der Archivar muß seinen Standort innerhalb der Schriftgutverwaltung, d.h. gegenüber dem Aktenproduzenten offensiver und selbstbewußter vertreten und verteidigen. Die Zusammenarbeit mit den Dokumentationsstellen, den Rechenzentren und den kommunalen Spitzenverbänden ist zu intensivieren.
6. Innerhalb der archivischen Diskussion muß der Stellenwert des kommunalen Sektors gestärkt werden. Dies gilt insbesondere für die Weiterentwicklung moderner Technologien, z.B. für Thesaurus-Entwicklung.
7. Dokumentarische Methoden und ADV-Anwendung müssen unter Archivaren mit größerer Zielstrebigkeit untersucht werden, der Erfahrungsaustausch der Anwender z.B. in einer Arbeitsgemeinschaft unabhängig von der archivischen Spartenzuordnung (ob kirchlich, staatlich, kommunal, kommerziell etc.) gefördert werden.

Ich komme zum Schluß: Obwohl die Rahmenbedingungen sowohl durch die finanzielle Situation der Archivträger als auch durch die häufig fehlende innovatorische Mentalität der Archivare selbst nicht gerade eine praktische Umsetzung neuer Verfahren nahelegt, hat sich also die ABSt entschlossen, das geschilderte System nunmehr als obligatorische Methode anzuwenden – zum erhofften Vorteil für Benutzer und Archivare.

Anhang: 10 Thesen zur gegenwärtigen archivischen Problemlage der Kommunalarchive im Landesteil Nordrhein.

1. Das durchschnittliche Kommunalarchiv gehört der unteren Größenklasse an und beinhaltet Schriftgut der preußischen Registratur ab 1815. Der Umfang des Archivgutes bis 1945 liegt bei etwa 3000 Einheiten. Nicht selten reichen Einzelbestände aber in die Frühe Neuzeit zurück. Jene Teile sind in ihrer Mehrheit repertorisiert.

Die Menge des neu produzierten Schriftgutes verdoppelt sich in ihrem Umfang in immer kürzeren Zeitabständen. Die Magazinkapazität kann nicht beliebig erweitert oder durch – mit hohen Investitionen verbundene – Einrichtungen aufgefangen werden. Durch die kommunale Neugliederung (ab 1969) verschärfte sich die Situation: Die Anzahl der Aktenproduzenten (und Aufbewahrer auch alter Archivalien!) verringert sich von 805 auf 165. Die andersartigen Probleme der 13 Kreisarchive bleiben hier unberücksichtigt.

2. Die Art des Zuwachses hat sich qualitativ verändert. Neuübertragungen und Umgestaltungen von Kompetenzen, erhöhte Beteiligungsfunktionen innerhalb und außerhalb der Verwaltung, schnellere und häufigere Strukturwandlungen etc. tragen zur Kompliziertheit der Arbeitsabläufe und damit der zugehörigen Akten bei. Sie sind nur noch für Spezialisten durchschaubar und verständlich. Dies gilt im Kommunalbereich auch deshalb verstärkt, weil die formale, die physische Aktenbildung die Sachverhalte kompakter und dichter zeigen muß als im Ministerialbereich mit weitläufigen Gliederungen. Durch die Vermehrung der parlamentarischen Gremien/Ausschüsse entstand völlig neues Schriftgut und verschiedene Arten der Doppelüberlieferung.
3. Durch höchst komfortable und verbreitete Vervielfältigungstechniken wird die Verteilung von Unterlagen aller denkbar beteiligten Stellen erleichtert, was zum Anschwellen des Aktenumfangs beiträgt. Hierzu tritt die Flut des Graubereichs (Drucke, Hefte).

4. Auf die tatsächliche Aktenbildung wird immer weniger Sorgfalt verwendet. Gliedernde und strukturierende Notwendigkeiten werden als "Formalismus" und Bürokratismus abgetan, d.h. zunehmend werden die Betreffe auf den Schriftstücken nach Gutdünken gebildet, Aktenzeichen weggelassen, mehrere Themen in einem Brief behandelt, Schreiben von vorne nach hinten oder umgekehrt nach Belieben abgeheftet, ein einheitlicher Aktentitel als überflüssig erachtet.

Die Akte, welche die Archivwissenschaft für ihre sogenannte "Titelaufnahme" voraussetzt, existiert nicht mehr.

5. Zur gleichen Zeit mußte der Archivar einen Wandel der Archivalien-Substanz hinnehmen. Speziell das Archivgut des 19. und 20. Jahrhunderts zeigt in der Substanz Auflösungserscheinungen durch hohen Holzschliffanteil im Papier, durch kurzfristig haltbare Fotokopien, durch rostende Ordner und Büroklammern, durch unsachgemäße Verwendung von Klebstoffen, durch aggressive Stempel-, Kopierstift- und Filzstiftfarben, neuerdings tritt das Recycling-Papier hinzu. In diesen Bereichen muß vielfach bei nicht-rettbaren Archivalien eine Zweitüberlieferung geschaffen werden oder der Archivar verzeichnet in Absprache mit dem Restaurator im Wettlauf mit der Uhr. Ein nicht geringes Problem (bei kleineren Archiven speziell) bringen ADV-Outputs auf maschinenlesbaren Datenträgern.

6. Der erhöhte Informationsbedarf der Parlamentarier und der Verwaltungsspitze erfordert schnelleren Aktenumlauf (d.h. lagerungs- und findtechnische Hilfen) sowie eine höchst mögliche inhaltliche Erschließung.

7. Dem damit angesprochenen erhöhten Ausstattungsgrad der Archive an Personal und sächlicher Einrichtung wird nicht entsprochen und ist angesichts der Finanzkrise nicht zu erwarten.

8. Unabhängig von der tatsächlichen Stellensituation im Kommunalbereich ist die Ausbildung der Archivare noch zu wenig auf die geschilderten Belange ausgerichtet.

9. Wissenschaft und Öffentlichkeit stellen hinsichtlich der Archivauswertung (Publikationen, Ausstellungen) immer höhere Anforderungen auch an kleinere Kommunalarchive. Bei diesen Anforderungen wird nicht selten auf Querschnittsfragen reflektiert (z.B. Preis des Bundespräsidenten zum Thema "Nationalsozialismus im Alltag"), die durch herkömmliche Erschließungsmethoden und Strukturkriterien (Provenienzprinzip) nur unzureichend beantwortbar sind. Althergebrachte Vorstellungen von Kassationsprinzipien (z.B. Massenakten) bedürfen der Überprüfung. Ferner müssen archivarische Erschließungsmethoden dem zunehmend archivfremden Benutzer (z.B. Schülern) gerecht werden.

10. Neue Aufbewahrungsformen, wie sie die Mikroverfilmung entwickelt hat, bedürfen neuer Erschließungsmethoden.

ÜBERNAHME UND ERARBEITUNG VON ADV-PROGRAMMEN DURCH DAS RECHENZENTRUM DES LANDSCHAFTSVERBANDES RHEINLAND

von R. Riebschläger

1. Organisation der Datenverarbeitung (DV) in Nordrhein-Westfalen (NW)

Das Gesetz über die Organisation der Automatisierten Datenverarbeitung in NW vom 12.2.1974 legt fest, daß die unter seinen Geltungsbereich fallenden Institutionen dem Gesetz entsprechend zu verfahren haben, sobald sie sich der Automatisierten Datenverarbeitung (ADV) bedienen wollen.

Das Gesetz bestimmt unter anderem, welche Institutionen welche Rechenzentren beanspruchen können. Für den kommunalen Bereich sind dies die Kommunalen Datenverarbeitungszentralen (KDZ), zu denen auch die KDZ des Landschaftsverbandes Rheinland (LVR) gehört. Sie ist für alle ADV-Angelegenheiten des LVR zuständig und damit auch Ansprechpartner der Archivberatungsstelle Rheinland.

Für die Gemeinden und kreisfreien Städte in Nordrhein-Westfalen bestehen entsprechende KDZ's.

2. Die KDZ des LVR

Die KDZ ist als "Referat für Verwaltungsrationisierung/ADV" in die Zentralverwaltung des LVR eingegliedert. In ihr sind zur Zeit 57 Mitarbeiter beschäftigt:

– Leitung und Verwaltung	4
– Systemanalyse/Organisation	14
– Programmierung	23
– Technischer Betrieb	13
– Datenerfassung	3.

Die Gesamtkosten betragen 1981 ca. 8 Millionen DM, wovon ca. 45 % durch Erträge von kostenerstattenden Benutzern gedeckt sind. Die Kosten teilen sich auf in ca. 47 % Personalkosten, 39 % Maschinen- und Programm-Miete und 14 % sonstige Kosten.

Die KDZ ist mit DV-Anlagen der Firma IBM ausgestattet. Bevorzugte Programmiersprache für Neuentwicklungen ist PL1. Letztere Angaben sind von Bedeutung hinsichtlich des Austausches von Programmen zwischen Rechenzentren. Trotz Verwendung höherer Programmiersprachen, ist der Programmaustausch zwischen Rechenzentren, die mit Anlagen unterschiedlicher Hersteller und/oder unterschiedlicher Betriebsprogramme arbeiten, problematisch.

3. ADV-Projekte der Archivberatungsstelle Rheinland

3.1 Vorbemerkungen

Die DV-Aufgaben im Bereich der Archivberatungsstelle Rheinland haben überwiegend die Bearbeitung von Texten in zunächst weitgehend unstrukturierter Form zum Gegenstand. Dies ist neu gegenüber den "traditionellen" DV-Aufgaben der KDZ, wo es hauptsächlich um die Verarbeitung stark strukturierter Daten geht (z.B. Lohn- und Gehaltsabrechnung, Krankenhausrechnungen, Buchhaltung usw.). Auf diese Aufgaben sind aber Maschinen und insbesondere die Standard-Programme (Software) und damit auch die Erfahrungen und Ausbildung der Mitarbeiter ausgerichtet.

Für die Lösung der auf diese neue Form der Daten ausgerichteten Aufgabenstellungen mußten also zunächst Erfahrungen und Kenntnisse gesammelt werden. Aus dieser Sicht sind auch die im folgenden genannten Projekte zu sehen. Ergebnisse dieser Erfahrungen sind Aufgabenlösungen, die einerseits den "Datenverarbeitern" wirtschaftlich vertretbare Programmentwicklung, andererseits dem Benutzer der Archivberatungsstelle Rheinland akzeptable Ergebnisse ermöglichen.

3.2 Projekt Findbuch-Indizierung

Für die nachträgliche Indizierung von Findbüchern wurden Programme erstellt, die aus den eingegebenen Benennungen (z.B. Orts- oder Personennamen) mit ihren Signaturen die entsprechenden Indizes ausgeben.

Schon bei dieser relativ einfachen Aufgabenstellung zeigen sich die Begrenzungen z.B. der Standard-Sortier-Software. Diese sortiert nämlich nur Zeichenketten ohne Berücksichtigung irgendwelcher Bedeutungen.

Beispiel:

Die Signatur 3.5 (Zeichenkette 3.5 ␣ (␣ = blank = Leerzeichen)) würde hinter 3.11 sortiert werden, da die 1 auf der 3. Stelle der Signatur kleiner als 5 ist.

Will man aber die Standard-Software nutzen, muß man die Daten entsprechend aufbereitet erfassen (d.h. im Beispiel 3.5 als 3.05).

Dies ist für den "traditionellen" Datenverarbeiter alltägliche Selbstverständlichkeit, für den Anwender aus dem Archivbereich beiläufig nicht.

Trotz dieser Einschränkungen erscheint die ADV gerade hier als sehr effizientes Hilfsmittel, da große Datenmengen in sehr kurzer Zeit geordnet und ausgegeben werden können. Darüber hinaus bietet sie den hervorragenden Vorteil, verschiedene im ADV-System gespeicherte

Indizes beliebig verknüpfen zu können, und zwar ohne jeglichen zusätzlichen Aufwand mit Ausnahme der reinen Maschinenkosten.

3.3 Projekt Bildarchiv

Für die Erfassung von Bildarchiven (als Pilot-Projekt ein privates) wurde ein bestehendes Verfahren einer Institution übernommen. Allerdings mußten einige Programme neu erstellt und einige geändert werden, da das Verfahren auf einem anderen Maschinensystem erstellt wurde. Hier wurden die in 3.1 angesprochenen Probleme des Programmaustauschs deutlich, zumal die Originalprogramme in Assembler und nicht in einer höheren Programmiersprache geschrieben sind.

Neben diesen technischen "Erfahrungen" wurde jedoch ein weiterer sachlicher Schritt vollzogen. Die Daten werden in Form "annotierter Deskriptoren" gespeichert und verarbeitet. Das heißt, die zu speichernden Daten werden zusätzlich mit einer Notation versehen. Aus dieser geht sowohl für den Anwender als auch für das DV-System hervor, um welche Daten es sich handelt. Dies ist die Voraussetzung für die Speicherung und Verarbeitung unstrukturierter Daten.

3.4 Projekt COMIS

Zum Inhalt des COMIS ist auf den Vortrag von H.P. Neuheuser über die "Verzeichnung der Gemeindearchive im Landesteil Nordrhein" zu verweisen.

In dem zur Zeit in der Realisierung befindlichen Teil COMADES finden die bisher gewonnenen Erfahrungen und Kenntnisse ihren Niederschlag. Das System arbeitet also mit annotierten Deskriptoren, was eine äußerst große Flexibilität bietet. Neben der bestehenden Anwendung für die Findbücher im Rahmen der Schriftgutverzeichnung ist es ohne sehr hohen Aufwand möglich, entsprechende Anwendungen für z.B. Bildmaterial und dergleichen mehr, zu realisieren.

Es müssen "nur" entsprechende Notationen festgelegt und die Auswertungsprogramme erstellt werden. Die Datenerfassung und -speicherung erfolgt für alle Anwendungen gleich.

4. Schlußfolgerungen

Die Übernahme von Programmen ist problematisch, wenn Maschinen und/oder Betriebs-Software des anwendenden Rechenzentrums und des produzierenden Rechenzentrums nicht übereinstimmen. Der Einsatz höherer Programmiersprachen bietet noch die günstigsten Voraussetzungen.

Die Erstellung von Programmen für das angesprochene Anwendungsgebiet erfordert gegenüber der "traditionellen" stark strukturierten DV zusätzliche Erfahrungen und Kenntnisse.

Der Aufwand für die Erstellung eines entsprechenden Anwendungssystems mit seinen Programmen ist nicht unerheblich, die Verteilung dieser Kosten auf mehrere Benutzer/Rechenzentren daher angezeigt.

Diskussion

O p p e l erkundigt sich nach der Höhe der Kosten des Communal Information System (COMIS) für die Gemeinden.

N e u h e u s e r antwortet, die Kosten würden vom Landschaftsverband Rheinland getragen, weil er als Träger dieses Projekts auftrete. Diese Situation werde aber im Hinblick auf die heutige Finanzlage nicht immer erhalten bleiben. Es sei in der Zukunft durchaus möglich, daß man Gebühren erhebe. Wie hoch diese ausfallen, sei jetzt nicht zu sagen. Er fügt noch hinzu, daß der Austausch von Programmen zwischen den Verwaltungen die Möglichkeit biete, kostenlos ein Programm zu erwerben. Die Kosten für einen Computer (Terminal-Retrieval) und seine Benutzung (Retrieval) müsse die Gemeinde selber tragen.

K e t e l a r fragt, ob die Archivberatungsstelle Rheinland auch mit anderen Ländern Programme austausche und wie die Lage in anderen Ländern sei.

R i e b s c h l ä g e r kann hierzu keine verbindliche Stellungnahme abgeben.

F r a u A u e r b a c h erklärt, in Hessen bestehe ein ähnliches System. Das Land Hessen unterhalte verschiedene Datenverarbeitungsinstitute für die kommunalen Verbände und ein Zentrales Rechenzentrum in Wiesbaden, die alle kostenlos benutzt werden könnten. Dies solle aber geändert werden. Weil alle Rechenzentren die gleiche Ausstattung haben, könnten die Programme leicht ausgetauscht werden. Bei der Frage, ob das COMIS-Modell auf Hessen zu übertragen sei, müsse man berücksichtigen, daß die meisten Kommunalarchive schon im Staatsarchiv aufgenommen seien.

R o m e y k will wissen, ob auch ein Retrieval-System geplant sei, welchen Benutzerkreis man zu erreichen glaube und auf welche Weise der Benutzer recherchieren könne.

N e u h e u s e r antwortet, daß ein Retrieval-System geplant sei. Das System böte drei Suchmöglichkeiten: 1. nach dem Klassifikationssystem, 2. nach den Listen und Indizes, und 3. alphabetisch. Die Archiveigentümer, so sei geplant, könnten dann über die kommunalen Rechenzentren on-line oder off-line Zutritt zu ihren Datenbeständen bekommen.

R o m e y k betont die Kostenfrage. Ein on-line-System sei wegen der Standleitung, der Inanspruchnahme der Computerzeit der Programme und der Telefonverbindung usw. mit erheblichen Kosten verbunden. Dies im Gegensatz zum Computereinsatz in der eigenen Verwaltung.

R i e b s c h l ä g e r bejaht die hohen Kosten eines on-line-Retrievals.

K e t e l a r fragt sich, ob der Benutzer für seine Untersuchungen bezahlen solle oder müsse.

N e u h e u s e r meint, daß dies eine kulturpolitische Frage und abhängig davon sei, in welchem Maße man eine benutzerfreundliche Archivpolitik verfolgen möchte. Die Archive müßten aber zunächst vor allem darüber nachdenken, inwieweit ein Retrieval-System noch notwendig sei, wenn Indizes und Listen optimal vorhanden sind, und sie müßten in Erwägung ziehen, ob und wie stark eine Nachfrage seitens der Benutzer bestehe.

O p p e l will wissen, wie ein kleiner Computer, der mit hardware arbeite und in vielen Verwaltungen vorhanden sei, dem COMIS angeschlossen werden könne.

S c h o t e l erklärt dazu, aus eigener Erfahrung festgestellt zu haben, daß man diese Frage der gegenseitigen Adaption eigentlich schon vorher geklärt haben müsse. Für Archive reiche im allgemeinen ein kleiner Computer nicht, der nicht mit einem auswärtigen größeren Computersystem harmoniere.

R i e b s c h l ä g e r erläuterte, daß das technische Problem einer Koppelung von zwei Systemen, wie O p p e l es sich vorstelle, nicht immer gelöst werden könne. Bei der Anwendung bestimmter Systeme gebe es jedoch sogenannte Schnittstellen-Programme. Bei diesem Verfahren müßten die beiden Programme, die auf die Peripherie zugehen, neu geschrieben bzw. umgestaltet werden. Der Kostenaufwand in einem solchen Fall sei seines Erachtens vertretbar.

R o e d o e erklärt, daß die Kleincomputer immer besser und leistungsfähiger würden, sie seien bereits fähig, sich auch der höheren Programmiersprache zu bedienen und ihre Speicherkapazität sei enorm gestiegen.

R i e b s c h l ä g e r gab zu bedenken, daß diese Systeme um die 150.000 DM kosteten.

R o e d o e erwidert, daß von seinem Ministerium ein solches System mit einer Speicherkapazität von 35 M schon für 90.000 DM erworben worden sei.

K e t e l a r fragt, ob auch in anderen Bundesländern solche Systeme wie COMIS aufgebaut würden und ob es möglich sei, dieses System zu "exportieren".

N e u h e u s e r hat sich in anderen Bundesländern umgehört und erkundigt und konnte feststellen, daß das COMIS derzeit noch einmaliges Projekt sei.

F r e s e meint, daß durch die immer weiter ausgedehnte elektronische Datenverarbeitung der Archivar sich mehr und mehr auf das Gebiet der Auswertung begeben, die eigentlich dem forschenden Wissenschaftler zukomme, der somit entlastet werde. Der Archivar solle sich auf die einfache Erschließung beschränken, und nicht wissenschaftliche Fragestellungen mit eingeben. Hierbei spiele dann auch die Frage, wer der geistige Eigentümer intensiverer Erschließung sei, deren Ergebnisse jederzeit von jedem abgerufen werden könnten.

R o m e y k hält dem entgegen, daß der Archivar nur so seiner Aufgabe gerecht werde und er die weiteren Ausarbeitungen, die durch den Computer ermöglicht würden, als einen günstigen Nebeneffekt betrachten solle. Doch solle sich der Archivar, der von F r e s e angedeuteten Gefahren immer bewußt bleiben, daß er nämlich zuweilen über die reinen Erschließungsbedürfnisse hinausgehe.

R i c h t e r i n g begrüßt das COMIS-System und seine vielseitigen Verwendungsmöglichkeiten. Er wendet aber ein, daß die Anwendung nur für hauptamtlich besetzte Kommunalarchive ausführbar sei, weil der Mangel an ausgebildetem Personal und an Zeit es nicht anders zuließe. Die gedanklichen Vorarbeiten, die bei der Archivberatungsstelle geleistet seien, seien gleichwohl von großem Wert.

N e u h e u s e r erklärte dagegen, daß das COMIS-System in erster Linie für nicht-hauptamtlich besetzte Archive gedacht sei. Durch die Verwendung des standardisierten Belegbogens und demnächst eines dezentralen, transportablen Datenerfassungsgerätes werde es möglich, auch örtliche Hilfskräfte auszubilden. Da verfüge er

aber nicht über Erfahrungswerte, denn bis jetzt seien nur Archivare des gehobenen Dienstes damit beschäftigt und keine Hilfskräfte. Die Initiatoren seien aber auch durchaus lernwillig und hätten daher eine Arbeitsgemeinschaft von ca. 20 Personen der COMIS-Anwender ins Leben gerufen. Dort diskutiere man die Kriterien und Voraussetzungen. Durch diese Arbeitsgemeinschaft und durch die intensive Betreuung seitens der Archivberatungsstelle sei es möglich, mehrere Projekte parallel zueinander zu bearbeiten, wie es zur Zeit mit den vier Pilotprojekten geschehe.

F r e s e weist darauf hin, daß die durch die Archivberatungsstelle angebotenen recht intensiven Dienstleistungen die Kommunalverwaltung dazu veranlassen könne, auf die Besetzung einer Archivberatungsstelle zu verzichten und die Last weiterhin von der Archivberatungsstelle getragen werden müsse.

Nach dieser Diskussion kam D r. M a n f r e d P e t r y (Hauptstaatsarchiv Düsseldorf) mit folgendem Referat zu Wort.

ERFAHRUNGEN BEIM EINSATZ DER EDV IM ARCHIV

von Manfred Petry

1. Voraussetzungen

1.1 Das Nordrhein-Westfälische Hauptstaatsarchiv (HStA) Düsseldorf benutzt die IBM-Rechenanlage des benachbarten Landesamtes für Datenverarbeitung und Statistik (LDS).

1.2 Geräte:

Datensichtgerät (DSG) oder Terminal – ein Bildschirm, an dem mit Hilfe einer angeschlossenen Schreibmaschinenähnlichen Tastatur unmittelbar Daten eingegeben und verändert werden können.

Am DSG werden ferner die Verarbeitungsläufe in Gang gesetzt. Seit 1978 in Betrieb.

Schreibmaschinen mit OCR-Typen bzw. Maschinen, die mit einem entsprechenden Kugelkopf ausgerüstet werden können (OCR – Optical Character Reader).

Formulartrenner oder Separator zum Trennen der Endlosformulare.

Nach der Trennung können die Ausdrücke in der Buchbinderwerkstatt des HStA zu dem gewünschten Format (meist DIN A 4) zurechtgeschnitten werden.

1.3 Programm

Die Index-Programme, um die es hier geht, wurden seit 1974/75 im HStA schrittweise entwickelt. Das IBM-System stellt sozusagen als Vorleistung lediglich die Sortierprogramme zur Verfügung.

1.4 Anwendungsmöglichkeiten

Indices zu bereits geschriebenen Findbüchern, Signaturen- und Seitenkonkordanzen mit Kennzeichnung von Akten, die über die übliche 30-Jahres-Frist hinaus gesperrt sind.

Darüber hinaus kann jedwede Art von Schriftgut durch Indices erfaßt werden. Dabei ist es an sich gleichgültig, ob das Schriftgut handschriftlich, maschinenschriftlich oder im Druck vorliegt. Im HStA wird inzwischen auch Bildmaterial mit den vorhandenen Programmen indexiert.

1.5 Kosten

Exakte Zahlen zu nennen ist schwer und zudem wenig hilfreich im Sinne einer verwertbaren Information, da die EDV-Arbeiten des HStA vom KDS bzw. von dem einschlägigen Titel mitgetragen und nicht gesondert angerechnet werden.

Die Jahresmiete für den Terminal beträgt derzeit rund 3.300 DM, also 275 DM im Monat. Die CPU-Minute, d.h. eine Minute reiner Rechenzeit kostet ca. 47 DM. In einem Monat nehmen wir ca. 50 CPU-Minuten, auf jeden Fall weniger als eine

Stunde reiner Rechenzeit in Anspruch. Der Drucker (Writer) bringt etwa 1000 Zeilen für 1 DM.

Der Belegleser kostet etwa 3 DM pro Minute; er wird von uns im Monat kaum mehr als eine halbe Stunde benötigt. Hinzuzurechnen wären diverse Sachkosten (z.B. Papier) und anteilige Personalkosten für das im LDS arbeitende EDV-Personal.

Soweit gesehen sind wir mit den im Hause laufenden EDV-Arbeiten pro Jahr etwa mit 75.000 bis 80.000 DM zu veranschlagen.

Nochmals die gleiche Summe entsteht dem HStA selbst an Personalkosten. Darin enthalten ist das Brutto-Gehalt für den im Hause arbeitenden EDV-Fachmann sowie das Gehalt einer Schreibkraft, die jedoch mit EDV-bezogenen Schreibarbeiten nicht das ganze Jahr ausgelastet ist.

Die hier theoretisch errechneten Gesamtkosten entsprechen in etwa den Bezügen von fünf Schreibkräften (BAT VII), doch ist dieser Vergleich wenig aussagekräftig, da die Einsatzmöglichkeiten der EDV ganz andere Perspektiven ermöglichen als die manuelle Tätigkeit von fünf Schreibkräften. Ich verweise hierzu auf die Ausführungen von Herrn Romeyk.

2. Organisation

2.1 Arbeitsgänge

2.1.1 Textgrundlage der Indexierung sind – wie bereits kurz erwähnt – maschinenschriftliche Findbücher oder gedruckte Texte. Das Einbeziehen handschriftlicher Texte ist prinzipiell möglich.

2.1.2 Die Erfassung erfolgt entweder mit Hilfe sogenannter Schreibmaschinen auf Endlosformular oder unmittelbar am DSG.

Bei maschinenschriftlicher Erfassung werden die zuvor getrennten Belegbögen in den Belegleser gegeben und in den Rechner eingelesen. Dieser Arbeitsschritt entfällt bei Erfassung am DSG.

2.1.3 Korrekturen

Nach dem Einlesen in den Rechner kann mit Hilfe einer einfachen Prozedur ein erster Ausdruck hergestellt werden, der dann sozusagen als Korrekturabzug dient.

Es empfiehlt sich, vor dem Bestellen dieses Korrekturabzugs bereits gewisse Fehler zu korrigieren, die erfahrungsgemäß immer wieder vorkommen und die sich ohne weiteres durch einfache Suchbefehle am DSG ermitteln lassen.

Solche Fehler sind etwa vom Rechner nicht erkannte Zeichen, die er in unserm System als \$ ausdrückt, oder fehlende oder falsche Code-Zahlen.

Liegt der Ausdruck vor, wird er vom Bearbeiter korrigiert. Die Fehler werden unter Verwendung der üblichen Korrekturzeichen gekennzeichnet und verbessert und anschließend am DSG eingegeben.

2.1.4 Ausdrucke

Alle gewünschten Ausdrucke lassen sich sowohl hinsichtlich der Sortierung wie hinsichtlich der Auswahl als auch, was die Anzahl betrifft, aus der zwar korrigierten, an sich aber nur einmal geschriebenen Eingabe (Erfassung) ableiten.

Wir pflegen bei normalen Findbuch-Indices zuerst die Konkordanzen auszudrucken, da sie Fehler wie etwa Doppel- oder Fehlnummern erkennen lassen. Es hat sich weiterhin als hilfreich erwiesen, mehrere Indices aus demselben Ressort miteinander in kumulierter Form ausdrucken zu lassen, da hierbei zuvor etwa übersehene Schreibfehler sozusagen zwangsläufig zutage treten, indem sie die vorgesehene Ordnung stören.

Eine nochmalige Trennung in Einzelindices ist damit keineswegs präjudiziert.

2.2 Personal

2.2.1 Schreibkraft

Die Schreibkraft braucht über die für ihre Tätigkeit ansonsten notwendigen Anforderungen hinaus keine weiteren Qualifikationen. Die im Umgang mit der EDV von ihr auszuführenden Tätigkeiten und Handgriffe lassen sich relativ rasch erlernen. Dies gilt auch für die Bewältigung einfacher Prozeduren wie etwa das Abschicken eines Druckjobs oder das Kopieren von Dateien aus dem Speicherbereich in den Arbeitsbereich, in dem die Korrekturen durchzuführen sind.

2.2.2 Der eigentliche Bearbeiter, d.h. der fachlich ausgebildete Archivar, wird normalerweise das Findbuch (oder eine Kopie davon) redaktionell so überarbeiten bzw. vorbereiten, daß die nunmehr rein manuelle Tätigkeit der Erfassung durch eine Schreibkraft erledigt werden kann.

Der Einsatz der Schreibkraft kann sich auch etwa auf das Heraussuchen der Personennamen oder gar nur auf die Übertragung der vom Bearbeiter kenntlich gemachten Korrekturen ins DSG beschränken.

Schließlich kann der Bearbeiter auch, falls die Umstände dies ratsam erscheinen lassen, alle Arbeitsgänge von der Erfassung bis zur letzten Korrektur selbst übernehmen. Dadurch würde ein Arbeitsgang und zugleich die damit verbundene zusätzliche Fehlerquelle eingespart bzw. ausge-

schaltet. Voraussetzungen dafür wären jedoch, daß einmal der Bearbeiter einigermaßen Übung im Maschinenschreiben hat und daß zum ändern die Schreibkraft weiterhin ausgelastet bleibt.

Es gibt im übrigen, wie noch auszuführen sein wird, Index-Arbeiten, für die normalerweise keine Schreibkraft zur Verfügung steht oder bei denen ihr Einsatz aus verschiedenen Gründen wenig rationell wäre.

2.2.3 Die Verarbeitungsläufe einer fertig korrigierten Datei besorgt der EDV-Fachmann im Hause. Es ist jedoch geplant, die Bearbeiter künftig auch in diesen Bereich schrittweise einzuführen. Voraussetzung dafür wäre zunächst einmal die Erweiterung der Anlage im Hause z.B. durch Aufstellen weiterer DSG. Die bisherige Konzentration aller Verarbeitungsläufe auf einen einzigen Mitarbeiter hat schon des öfteren zwangsläufig zu empfindlich spürbaren Engpässen, d.h. zu Wartezeiten geführt.

3. Bisherige Projekte

3.1 Ich selbst arbeite seit etwa zwei Jahren an maschinell verarbeiteten Indices. Als erstes hatte ich seinerzeit einen Index aller erreichbaren Organisations- bzw. Geschäftsverteilungspläne des von mir als Archivar zu betreuenden NW Kultusministeriums angefertigt. Er dient als dienstliches Arbeitsmittel, d.h. als Handexemplar des zuständigen Dezernenten

1. als Grundlage einer Behördengeschichte
2. als schnell überschaubare Tabelle der Wanderrichtung bestimmter Ressorts
3. als Hinweis auf die bislang noch nicht ausreichend im Hause dokumentierten Bereiche des Ministeriums, auf die der Behördenkontakt sich folglich zu konzentrieren hat
4. als eine Art Thesaurus, der teils im Gedächtnis, teils nachschlagbar den bei Verzeichnungsarbeiten notwendigen Wortschatz enthält.

3.2 Findbuch-Indices und Konkordanzen

Als nächstes wurden von meiner Sachbearbeiterin und mir rund 30 Findbücher der bisher verzeichneten Bestände des Kultusministeriums indexmäßig erfaßt, d.h. aus der einmal geschriebenen Eingabe wurden

1. die Einzelindices
2. ein Sammelindex
3. ein Personenindex
4. Konkordanzen mit Angabe der Sperrfristen abgeleitet.

3.3 Anwendung auf mittelalterliches Schriftgut

Es lag nahe, einen Versuch zu unternehmen, die bei der Indexierung von modernem Aktenschriftgut gewonnenen Erfahrungen auch einmal auf älteres Quellenmaterial anzuwenden.

3.3.1 Projekte

Objekt dieses Versuches war das Kettenbuch des Stiftes Essen, ein um 1410 niedergeschriebenes Urbar, das die Namen, Besitzer und Abgaben von über 1000 Höfen nennt. Aus der einmal erfaßten Eingabe wurden

1. ein Ortsindex
2. ein Personenindex
3. ein Abgabenindex

in jeweils mehreren Sortierungen generiert.

Aus der gleichen Eingabe wurde eine tabellarische Übersicht der wichtigsten bzw. häufigsten Abgabenarten (Getreidesorten, Vieh und Geld) abgeleitet, die in einer Art Diagramm Zusammenhänge erkennen läßt, die bei der Lektüre der Quelle oder bei Zuhilfenahme der Indices nicht so schnell ins Auge springen. (Die dabei erkannten Zusammenhänge sind natürlich noch nicht das eigentliche und gewünschte Forschungsergebnis, sondern sie sind als ein Denkanstoß zu betrachten, in eine bestimmte Richtung weiterzuforschen.)

Desgleichen wurden rund 400 Urkundenregesten indexiert.

Dieser Index wurde einmal in herkömmlicher Weise wie die modernen Findbuch-Indices ausgedruckt und dann mit relativ geringem Zeitaufwand für einen zweiseitigen Ausdruck in Groß- und Kleinschreibung mit Silbentrennung und Zeilenausgleich umgewandelt.

3.3.2 Probleme und ihre Lösung

Mittelalterliche Quellen bieten eine Reihe von Lauten, für die weder auf der Schreibmaschine noch auf der Tastatur des DSG eine Type verfügbar ist, z.B. ũ oder ö.

Es gibt ferner Buchstaben, die nicht ihrem graphischen Aussehen, sondern ihrem Lautwerk entsprechend sortiert werden müssen, z.B. vokalisches v oder konsonantisches u.

Zur Behebung der erstgenannten Schwierigkeit stehen im IBM-System zwar Programme zur Verfügung, mit deren Hilfe solche Buchstaben hergestellt werden können, doch haben wir uns diese Programme aus Zeitmangel bislang noch nicht zu eigen gemacht. Ich greife derzeit im Ausdruck noch zu einer Hilfskonstruktion, z.B. [uo] für ũ.

Die zweite Kategorie von Buchstaben, in die die erstgenannte selbstverständlich einzubeziehen ist, wird bei der Erfassung als Zahl geschrieben, z.B. konsonantisches u als 9. Über eine Steuerdatei,

in der alle verwendeten Chiffren aufgelistet sind, wird der Rechner im Zuge der Verarbeitungsläufe dann veranlaßt, diese 9 zwar als u auszudrucken, sie jedoch wie v zu sortieren.

Auf vergleichbare Weise, nämlich durch das Hinzufügen einer Steuerzahl, läßt sich schließlich ein weiteres Problem lösen, das jedem vertraut ist, der schon einmal einen Index zu einem mittelalterlichen Quellenwerk angefertigt hat. Gemeint ist das hin und wieder notwendige Durchbrechen der strengen alphabetischen Folge der Sortierung durch eine hierarchische oder logische Gliederung. Mit dieser Technik werden auch Namens- und Sachindex bei der Sortierung auseinanderdividiert.

Vor dem Ausdruck werden diese Steuerzahlen per Programm wieder eliminiert.

Für alle hier geschilderten Prozeduren bei der Indexierung mittelalterlicher Quellen konnte auf die Programme zurückgegriffen werden, die zuvor für die Indexierung modernen Schriftguts erarbeitet worden waren.

4. Die Ergebnisse der bisherigen Indexarbeiten und die praktischen Anwendungsgebiete wurden in den voraufgehenden Ausführungen schon beiläufig genannt, so daß wir uns hier auf eine knappe Zusammenfassung beschränken können.

4.1 Bestimmte Findhilfsmittel, wie etwa der Index der Organisations- und Geschäftsverteilungspläne sind ausschließlich für die Arbeit auf dem Dienstzimmer bestimmt. Er kommt dem Benutzer nur insoweit zugute, als er zur umfassenden und rationellen Beratung durch den Dezernenten oder Sachbearbeiter herangezogen werden kann.

Aus Gründen des Persönlichkeits- und des Datenschutzes gilt das gleiche für die kumulierten Personenindices zu modernem Schriftgut.

Konkordanzen dienen dem zuständigen Archivar ebenso wie der Benutzersaalaufsicht als schnelle Orientierungshilfe in der Frage, ob eine bestimmte Akte vorgelegt werden darf oder nicht und auf welcher Seite welchen Findbuches die genaue Titelaufnahme zu finden ist.

4.2 Die übrigen Einzel- und Sammelindices sind auch dem Benutzer zugänglich. Das gilt selbstverständlich auch für den Index zu den Urkundenregesten.

4.3 Die Indices zu dem genannten Urbar, dem Kettenbuch, sind schon eher für die wissenschaftliche Forschung konzipiert, wenngleich sie auch zur Benutzerberatung dienlich sein können.

5. Ausblick

5.1 Auf die vielfältigen Möglichkeiten, die sich aus der maschinellen Verarbeitung von Indices ergeben, hat Herr Romeyk schon in gebührender Breite hingewiesen.

- 5.2 Die Schwierigkeiten der Bearbeitung mittelalterlicher Quellen können zwar hin und wieder Kopferbrechen bereiten, vielleicht sogar heftige Kontroversen auslösen. Sind sie aber einmal im konventionellen Bereich gelöst, so bereitet ihre maschinelle Bewältigung keine grundsätzlichen Probleme mehr.
- 5.3 Der Einsatz der EDV gibt uns vielmehr die Möglichkeit, in Bereiche vorzustößen, an die man zunächst gar nicht denkt oder die erst gar nicht in Angriff genommen werden, weil sie einer Art Sisyphusarbeit gleichkommen.
- So würde man für das Notieren gelegentlicher Notizen, die man in eine langsam wachsende Kartei oder einen Zettelkasten steckt, niemals eine Schreibkraft oder auch nur eine Schreibmaschine bemühen. Ich habe mir inzwischen zur Gewohnheit gemacht, solche Notizen gleich nach dem vertrauten Indexmodell auf DIN-A 4-Bögen aufzuschreiben und diese Listen dann von Zeit zu Zeit am DSG einzugeben.
- 5.4 Mit dem vorgestellten Indexmodell könnten z.B. Urkundenbestände, deren mühsame Regestierung Jahre und Jahrzehnte in Anspruch nehmen würde, auf Orts- und Personennamen sowie auf wichtige Sachbegriffe abgesucht werden. Dies kann unmittelbar am Original geschehen. Die Regestierung oder gar Publizierung soll damit nicht überflüssig gemacht werden. Aber zunächst gilt es doch, auch für die älteren Archivbestände im Hinblick auf die ständig steigenden Benutzeransprüche gewappnet zu sein. Diesen Anforderungen sind die alten Archivverzeichnisse doch kaum noch gewachsen, weil sie für ganz andere Zwecke und Fragestellungen konzipiert wurden.
- 5.5 In vielen Archiven, um nicht zu sagen in den meisten, liegen noch ganze Serien von Rechnungen und Protokollen unverzeichnet, andere Gruppen wie etwa Prozeßakten und Korrespondenzen sind vielfach nur summarisch erschlossen. Vielleicht bietet die neue Technik, die ja mit einer nicht unbeträchtlichen Einsparung von Schreibarbeit verbunden ist, einen neuen Anreiz, auch solche Quellengattungen zugänglich zu machen.

Fragen nach dem Vortrag von Herrn Petry

F r e s e fragt, ob Petry sich selbst hinter die OCR-Maschine gesetzt und die Endlostexte selbst eingegeben habe und dann weiter, ob Wortvariationen nicht in unterschiedlicher Typographie ausgedruckt werden könnten, wie es beim konventionellen Druckverfahren möglich sei.

P e t r y bestätigt zunächst, daß er die OCR-Maschine selbst bedient habe und erklärt dann, der von ihm erstellte Index sei ein Teil eines wachsenden Gesamtindexes, und daher träten schon Probleme bei der Verwendung verschiedener Schriften auf. Außerdem bedeute das Heranziehen weiterer Schrifttypen für den Computer mutatis mutandis ungefähr die gleiche Arbeit, die eine Druckerei für den Entwurf neuer Buchstaben brauche.

F r e s e meint, feststellen zu können, daß heutzutage mehr nach Personen und Orten, weniger nach Sachthemen gefragt werde.

P e t r y glaubt, die Ursache sei darin zu sehen, daß ein Großteil der Besucher aus Genealogen und Heimatforschern bestehe. Die kleinere Zahl der wissenschaftlich vorgebildeten Forscher finde sich auch in weniger gut aufgearbeiteten Findbüchern noch gut zurecht. Es sei auch festzustellen, daß Sachindices selbst heute noch stiefmütterlich behandelt würden, sie seien überhaupt recht selten.

F r e s e fragt, wie die Wörter mit verwandten Konsonanten wie "B" und "P" oder mit Vokalen miteinander in Verbindung gebracht werden.

P e t r y antwortet, dieses Problem stelle sich auch bei der manuellen Bearbeitung eines Index. Man versuche dies durch Verweise zu lösen, z.B.: Oldendorpe siehe: Altendorf; Altendorf (Aldendorpe), doch arbeite und sortiere ein Computer dabei rationeller.

O p p e l fragt, ob der Thesaurus, den Dr. Petry benütze, auch außerhalb des Hauptstaatsarchivs zu erhalten sei.

P e t r y bedauert, dieser Thesaurus stehe nur für den dienstlichen Gebrauch zur Verfügung.

Darauf schlägt *O p p e l* vor, den Thesaurus durch Publikation einem größeren Kreis zur Verfügung zu stellen.

P e t r y hält diesen Vorschlag für erwägenswert und sagt, daß man zur Zeit einen Thesaurus für den gesamten Bereich des Ministerialarchivs plane.

Nach diesen Fragen übergaben die Moderatoren das Wort an H e r r n J. S p o o r (Samsom Data-processing, Alphen a.d. Rijn) *als letzten Referenten der Sitzung.*

DIE ANWENDUNG UND MÖGLICHKEITEN DER EDV BEI DER ERSCHLIESSUNG VON ARCHIVGUT

von J. Spoor

Obwohl ich kein Spezialist in Sachen Archiv bin, meine ich dennoch zur Gedankengestaltung um die EDV-Anwendung im Archivbereich beitragen zu können. Diese meine Meinung basiert auf einer langjährigen Erfahrung im Gebrauch von EDV-Mitteln zur Erschließung verschiedenartiger Dokumentensammlungen.

Seit 1970 habe ich mich mit einer großen Zahl von Projekten, die die Publikation von Dokumentensammlungen betreffen, beschäftigt. Diese Projekte durchlaufen einige Stadien, bevor die Publikationsentscheidung getroffen wird. Und diese Stadien sind es, die das Thema dieses Vortrags bilden, denn ich sehe hierin Anknüpfungspunkte zu identischen Stadien in der Erschließung von Archivgut. Zur Kennzeichnung der verschiedenen Möglichkeiten gehen wir aus von einer Dokumentensammlung, mit der Nachfolgendes gemacht werden könnte:

1. Die vollständige Publikation der Dokumente mit Registern,
2. die gekürzte Publikation,
3. das Zugänglichmachen dieser Dokumentensammlungen über Online.

1. Die vollständige Publikation der Dokumente mit Registern

Im traditionellen Verfahren müssen dabei die Register mit der Hand angefertigt werden. Das ist ein arbeitsintensives Verfahren, und ich nehme an, daß Sie mit dieser Weise des Publizierens bekannt sind. Moderne, vom Computer gesteuerte Photosatztechniken im graphischen Betrieb machen es möglich, dieses Publikationsverfahren auf andere Art zu regeln. Besonders die Satzkosten im graphischen Betrieb können bedeutend gesenkt werden. Man kann das arbeitsintensive Verfahren zur Registeranfertigung vermeiden. Dies wird erreicht dadurch, daß man die Eingabephase nicht länger im graphischen Betrieb verarbeiten läßt, sondern die Möglichkeiten der EDV verwendet. Die zu publizierenden Dokumente müssen dazu einer Aufbereitung unterzogen werden. Wörter oder Begriffe, die im Register erscheinen sollen, müssen kodiert werden. Soll die Publikation der Dokumente in einer bestimmten Reihenfolge stattfinden, muß dies ebenfalls in die Dokumentenbeschreibung aufgenommen werden. (In dieser Weise ist es auch möglich, dasselbe Dokument an mehreren Stellen in der Publikation aufzunehmen.) Der folgende Schritt ist die maschinenlesbare Fertigung der Dokumententexte und der hinzugefügten Informationen.

Hierzu gibt es verschiedene Möglichkeiten:

- Datenstapelung mittels eines Textverarbeitungsapparates (Wordprocessor)
- Datenstapelung mittels eines "Personal Computers"
- Ausführung der Aufgabe durch ein Computerservicebüro.

Eine sehr genaue Prüfung des gestapelten Textes und der Kodifizierung soll nachdrücklich beachtet werden. Nach der Eingabe wird der Text in den externen Datenspeicher eingetragen. Mit Hilfe der dazu entwickelten Computerprogramme können die Dokumente nach jeder gewünschten Reihenfolge sortiert und gleichzeitig Register angefertigt werden. Der genaue Zeitpunkt der Registeranfertigung wird dabei von der gewählten Einordnungsmethode bestimmt. Wählt man eine Verweisung nach Dokumentennummern, so findet die Registeranfertigung aus dem externen Datenspeicher statt. Wenn es notwendig ist, nach Blattnummern zu verweisen, so hat zuerst der graphische Prozeß einschließlich des Blatt-lay-outs stattzufinden, bevor die Register erarbeitet werden können.

Nach diesem Schritt erfolgt die Übertragung des Textes auf ein anderes Medium, das als Photosatz zu verwenden ist. Alle das Photosatzgerät steuernde Kommandos müssen im gespeicherten Text vorhanden sein. Wenn es sich um Dokumententexte handelt, werden die Kommandos dem Text während der Eingabephase zugefügt. Hinsichtlich der Register leistet das Computerprogramm diese Zufügungen. Der graphische Betrieb hat jetzt den Photosatz und kann die graphische Verarbeitung durchführen.

Der gespeicherte Text kann für verschiedene Zwecke benutzt werden, zum Beispiel für

- die Datenausgabe auf Microfiche (COM)
- das Durchführen von Teillektionen
- das Kombinieren der Dokumentensammlung mit anderen bereits zur Verfügung stehenden Sammlungen
- die Erschließung der Dokumente mittels Online-Verfahren.

Das Corpus Gijsseling¹ mit Dokumenten aus der Zeit vor 1300 wurde durch uns auf diese EDV-gestützte Weise publiziert. Dabei entstanden ebenfalls ein retrograder Index aller mittelniederländischen Wörter sowie ein Wortfrequenz-Index und ein Wortlängen-Index.

2. Die gekürzte Publikation

In dieser Publikationsform erscheinen nur Informationen über die Dokumentensammlung und die darin vorhandenen Einzeldokumente oder Dokumentengruppen. Der Typus des klassischen Niederländischen Archivinventars mag hierfür als gutes Beispiel gelten. Jedes Dokument oder jede Dokumentgruppe ist so nummeriert worden, wie sie nach der Titelerfassung in bestimmter Reihenfolge eingeordnet wurden. Je Dokument oder Dokumentgruppe fallen die nachfolgenden Informationen an:

- die Dokumentenart (Urkunde, Brief etc.);
- Betreff des Dokuments;
- die Namen der Interessenten und Ortsbezeichnungen;
- das Datum, materielle Form und Umfang der Einheit.

Die Anwendung der EDV zur Inventaranfertigung hätte nach folgender Phasierung stattfinden können:

A. Jedes Dokument wird in ein Formular eingetragen.

Dieses Formular kennt einen Entwurf, der es dem Dokumentenerfasser ermöglicht, sich nur mit den für ihn wichtigen Sachen zu beschäftigen. Ein derartiges Formular hat besondere Felder für die folgenden Daten:

- die (Unter-) Abteilung, in die das Dokument eingetragen werden soll (wenn nötig, kann das fragliche Dokument auch an mehreren Stellen eingetragen werden);
- die je Register gewünschte Erwähnung (Stichwörter oder Thesaurustermini);
- die Quellenangabe des Dokuments;
- die einfache oder mehrfache Klassifizierung (zum Beispiel U.D.C., alte Registraturnummern . . .);
- das Datum des Dokuments;
- der Betreff des Dokuments;
- der gekürzte Inhalt des Dokuments [etc., etc.].

B. Dieses Verfahren hat folgende Vorteile:

- es ist ein Hilfsmittel für eine schnelle und vollständige Beschreibung des Originaldokuments;
- der gekürzte Dokumenteninhalte bedarf nur einer einmaligen Erwähnung;
- es bietet Möglichkeiten zur Standardisierung, wodurch Kombinationen mit gleichartigen Sammlungen erreicht werden können;
- es bietet Möglichkeiten, mit Hilfe der EDV die Publikation und die Register zu fertigen, indem der Computer die folgenden Handlungen verrichtet (nachdem die Daten maschinenlesbar angefertigt worden sind):
 - Sortierung des Originaltextes nach gewünschter Reihenfolge;
 - Eintragung des Originaltextes an mehreren Stellen im Buch;

- Numerierung;
- Selektierung der Registereintragungen;
- Sortierung der Register;
- Entfernung von identischen Registereingängen;
- Zufügung der Satzkommandos am Hauptband und Übersetzung auf ein Photosatzmedium.

C. Die eingetragenen Originalformulare stehen immer zu weiteren Anwendungen zur Verfügung, z.B.

- zur vollständigen oder teilweisen Zusammenfügung mit anderen gleichartigen Datensammlungen;
- zu verschiedenen Indizes, die nicht in die Publikation aufgenommen werden, wie:
 - Sachindex
 - Konkordanz zwischen Aktenzeichen und Signatur
 - Online-Speicherung für spätere Beratung, zum Beispiel für wissenschaftliche Zwecke
 - Datenspeicherung für spätere mehr ausgedehnte oder geänderte Neuauflagen, weil die Dokumente aktualisierbar sind.

3. Die Zugänglichkeit von Dokumentensammlungen über Online

Sowohl bei der vollständigen wie bei der gekürzten Publikation gehört, wie ich bereits erwähnte, auch die Online-Speicherung zu den Möglichkeiten. Außer der genügenden Online-Speicherkapazität braucht man Computerprogramme zum Aufbau einer derartigen Online-Speicherung, zur Eintragung der Daten sowie zur Datenberatung. Weiterhin muß das Rechenzentrum über Datenkommunikationsmöglichkeiten verfügen.

Ortsmäßig soll das Computerzentrum ebenfalls über verschiedene nationale Netzanschlüsse und/oder einen internationalen Netzanschluß (wie Euronet, Telenet usw.) verfügen können. Aus dem gekürzten oder vollständigen Text werden einige Computerindizes angefertigt, die das Originaldokument selbst und innerhalb des Originaldokumentes indizieren.

Zur Datenerfassung muß der Benutzer über ein Terminal verfügen. Eine Telefonanlage verbindet das Terminal mit dem Computersystem, und der Benutzer kann durch einfache Kommandos eine Suchaktion durchführen. Die eingetippten Suchindizes werden durch Boolean Operatoren zur Suchstrategie zusammengefaßt. Erbringt diese Suchaktion eine ausreichende Zahl von Dokumenten, so können diese vom Terminal ausgedruckt werden. Ausgedehnte Ausgaben werden vorzugsweise im Rechenzentrum ausgedruckt. Der Wert der erhaltenen Informationen für den Benutzer zeichnet sich aus durch

$$\text{Präzision} = \frac{\text{richtige gefundene Dokumente}}{\text{alle gefundenen Dokumente}}$$

Vollständigkeit = $\frac{\text{richtige gefundene Dokumente}}{\text{alle richtigen Dokumente, die in der Welt zu finden wären}}$

Der Benutzer empfindet die Vollständigkeit als eine wichtigere Sache als die Präzision.

Der Zweck der Informationserschließung ist es, die beiden Ziele soweit wie möglich befriedigend zu erreichen. Die Verhältniszahl sollte so nahe wie möglich bei eins liegen.

Erschließungsmethodik

Mit Erschließungsmethodik ist die Methode gemeint, in der die vorhandene Information in einer derartigen Weise gekennzeichnet wird, daß die in einem bestimmten Augenblick gewünschte Information schnell selektiert werden kann. Wir kennen:

1. äußere Kennzeichen wie:

- Quelle
- Name
- Dokumentenart
- Dokumentendatum
- Sprache.

2. inhaltliche Kennzeichen (Gegenstände, Sachbetreffe).

In der Regel bietet in technischer Hinsicht die Erschließung nach äußeren Kennzeichen wenig Schwierigkeiten. Dennoch soll im voraus eine gute Kennzeichenanalyse der zu verwendenden Kennzeichen durchgeführt werden, weil spätere Ausdehnung der bereits aufgebauten Datenbank sich als schwierig erweist. Die Erschließung nach inhaltlichen Kennzeichen (auch Gegenstandserschließung genannt) ergibt weitaus mehr Schwierigkeiten. Es gibt mehrere Methoden solcher Erschließung. Wir nennen:

1. Kennzeichnung durch Stichwörter ohne stringente Regeln, meistens auf dem Text des Originaldokuments basierend.
2. Die automatische Textanalyse, mit Anwendung der gefundenen Termini als Suchtermini.
3. Die Anwendung eines einfachen Schemas mit einer Anzahl von Gruppen, nach denen die Dokumente eingeordnet worden sind.
4. Die Anwendung einer Klassifizierung oder Taxonomy (d.h. ein im voraus gemachtes Einordnungsschema); gegebenenfalls einschließlich eines alphabetischen Registers.
5. Die Anwendung eines Thesaurus, das ist eine Liste mit alphabetisch eingeordneten Wörtern nach bestimmten Regeln und mit wechselseitigem Bezug.

Die Verwendung freier Stichwörter (1.) ist die billigste Lösung. Die Qualität eines derartigen Systems ist aber gering. Die automatische Textanalyse (2.) bedarf der Eingabe des ganzen Textes. Das ist, wenn eine Speicherung auf Magnetband oder einem ähnlichen Medium

nicht schon vorhanden ist, kostspielig. Wohl kann mit Hilfe dieser Technik die Fertigung einer Suchliste mit allen gespeicherten äußeren Kennzeichen erreicht werden (wie z.B.: Wörter in "Abstracts" und ähnliches). Man kann es als Hilfsmittel anwenden.

Die Probleme der Klassifizierung (3., 4.) zeigen sich bei der Ausdehnung der Gegenstandszahl oder bei der Änderung des Einordnungsschemas: alle bereits eingeordneten Dokumente müssen wieder erneut eingeordnet werden. Deshalb ist Klassifizierung nur in stabilen Themenbereichen angebracht.

Die elektronischen Systeme verwenden heutzutage in der Regel Thesauri (5.). Diese ergeben die besseren Möglichkeiten zur Erschließung und zum Retrieval der Dokumente. Aber der Aufbau eines Thesaurus ist teuer.

Als Preisindikation gilt etwa achtzig bis hundert Mark je Terminus (Term). Ein Thesaurus geringen Umfangs besteht aus ungefähr 3000 Terms, während ein umfangreicher Thesaurus 12 – 15.000 Terms kennt. Die Entscheidung für eine bestimmte Erschließungsmethode kann nicht eher getroffen werden, als genügend bekannt ist, um welche Themenbereiche es sich handelt und wie

viele Dokumente einzutragen sind. Manchmal sind bereits Klassifizierungen oder Thesauri vorhanden. In diesem Falle soll man ihrer Anwendung im Hinblick auf die Kosten den Vorzug geben.

Die ausgewählte Erschließungsmethode hat Folgen für die Präzision und die Vollständigkeit, die von vier fundamentalen Verbalkommunikationsproblemen beeinflusst werden:

1. Syntaktische Probleme

Die Reihenfolge zweier Wörter kann als bestimmend gelten für die Bezeichnung der Wörter.

Eine umgekehrte Reihenfolge kann etwas ganz anderes bedeuten (z.B. in der Methodologie oder in der Geschichte).

2. Semantische Probleme

z.B.:

- a. Synonyme wie "Mathematik" oder "Algebra" und wie "Telegramm" oder "Depesche".
- b. Homographen wie Schimmel (Pilz, Pferd).

3. Grobe oder verfeinerte Wörter

z.B.:

Heilkunde, Augenkrankheit, Star.

4. Auch der Gesichtspunkt kann von Bedeutung sein.

z.B.:

Alkoholprobleme:

- medizinisch
- sozial
- Bekämpfung
- Brennstoff
- steuertechnisch
- illegales Branntweinbrennen.

Je besser die Erschließungstechnik, desto höher die Kosten, aber auch: desto größer die Präzision und die Vollständigkeit.

Ich möchte hiermit meinen Beitrag beenden und hoffe, Ihnen einen Überblick über die Möglichkeiten der EDV und ihre Anwendung zur Erschließung von Dokumentensammlungen gegeben zu haben.

1 M. Gijsseling, Corpus van Middelnederlandse teksten (tot en met het jaar 1300). 1. Ambtelijke bescheiden, 's-Gravenhage 1979. Besprechungen z.B. in: Naamkunde (1979), S. 20 – 43.

Schlusdiskussion

Ketelaar eröffnet die Diskussion und fragt, ob es bei SAMSOM schon einen Thesaurus mit Deskriptoren für Kommunalverwaltungen gäbe.

Spoor verneint das. Wohl gäbe es Thesauri für den maritimen Bereich und für den Lebensmittelbereich.

Ketelaar erwähnt noch, daß die Bestandsübersicht des Stadtarchivs Amsterdam, die in einigen Tagen erscheinen werde, auch mittels Computer bei SAMSOM bearbeitet worden sei. Er frage sich aber, wie man Begriffe, die nicht ausdrücklich im Text vorkommen, in den Index miteinbeziehen könne.

Spoor antwortet, daß erst ein vorläufiger Index gemacht würde, in den dann die Redaktion die zusätzlich gewünschten Änderungen und Erweiterungen einbringen könne.

Romeyk erläutert, daß dies der Ansatzpunkt für Wortlisten sei, die in der Rückwirkung auf die Praxis die Entscheidung bringen, bestimmte Sachverhalte so gestalten, wie man sie mit vorgefundenen Begriffen bereits beschrieben habe. Er nennt als Beispiel einen Index für ältere Akten aus der Regional- und Kreisverwaltung, in denen verschiedene Bezeichnungen häufig gewechselt hätten, z.B. "Gesundheitsverwaltung" und "Medizinalverwaltung", dabei handle es sich materiell um die gleiche Sache, nur unter unterschiedlichen Benennungen. Man habe daher vorab einen Index aus den eingereichten Manuskripten erstellt. Der daraus entstandene Index wurde den Bearbeitern mit der Auflage zurückgegeben, die weitere Indizierung zu harmonisieren. Diese Begriffsliste bestand aus ungefähr 3000 Wörtern – eine Zahl, die sich durchaus noch reduzieren ließe – und sei zur Erarbeitung einer Deskriptorenliste hilfreich.

VandeVoort fragt danach, wieviel Zeit man brauche, um die Daten einer Aktenerfassung auf ein Formblatt zu schreiben.

Neuheuser schätzt den Zeitaufwand bei der Zeichnung neuzeitlicher Akten für COMIS auf durchschnittlich 20 Minuten. Diese Werte beziehen sich aber nur auf die ersten 1000 Beschreibungen.

Schootel kann für Dordrecht keine genauen Angaben machen. Die Akten seien zu unterschiedlich. Für die eine Akte brauche man zwei Stunden und für eine andere Akte zwei Tage, weil in den Notariatsakten bisweilen viele Erben aufgezählt würden oder das Erbgut viele Häuser usw. umfasse.

VandeVoort unterstreicht jedoch, daß es vor allem im Hinblick auf die Planung und die Kosten-Nutzenrechnung sehr wichtig sei zu wissen, wie lange man brauche.

Neuheuser fügt noch hinzu, daß bei dem Belegbogen (Formular) noch ein spezielles Steuerzeichen

eingesetzt werden könne, um vorher eingegebene Daten zu wiederholen. So könne man vor allem bei Serienakten, z.B. Haushaltsplänen, Personalakten etc. viel Zeit einsparen.

Schootel sagt dazu, er habe für sein Projekt kalkuliert, daß 1 1/2 Personen in zwei Jahren Arbeitszeit Notariatsakten eines Zeitraums von 10 Jahrgängen erschließen werden, mit anderen Worten ca. 8000 Akten in 3 Jahren bei einer Arbeitskraft.

Romeyk kommt auf das Formblatt zurück. Er habe beobachten können, daß bei dem Gebrauch von Formblättern als Ersatz für das klassische Verzeichniskärtchen eine gewisse Zurückhaltung bestehe. Es werde vorgebracht, daß die Kategorien der Formblätter zu sehr reglementierten, daß sie Raum verschwendeten, wo eine kurze Angabe ausreiche, umgekehrt aber auch manchmal zu wenig Raum böten. Er bekenne sich trotz dieser Einwände zu dem Formblatt. Es biete Vorteile bei der Bearbeitung, weil z.B. bestimmte Angaben immer an der gleichen Stelle eingetragen würden, das sei ein heilsamer Zwang zur Disziplinierung. Andererseits könne das Formblatt auch durchaus flexibel sein. Der Entwurf eines Formblattes sei ziemlich einfach, und die Multiplizierung könne durch ein einfaches Kopiergerät geschehen. Für die Schreibkräfte seien solcher Art erfaßte Daten einfacher zu verarbeiten.

Ketelaar weist auf die Möglichkeit hin, ein solches Formblatt auf den Bildschirm zu projizieren [eine Maske], oder Fragen und Gegenfragen in ein Programm zu verarbeiten.

Romeyk gibt zu bedenken, daß diese Art zu arbeiten eine komplizierte und teure Technik voraussetze, über die erst wenige Archive verfügten.

Oppel möchte geklärt wissen, wieviel Zeit benötigt werde, um ein solches Formblatt zu erstellen.

VandeVoort sagt darauf, daß er die in England entwickelten Formblätter übernehmen konnte, weil die Beschreibung von Museumsgegenständen universal ist. Er hätte also keine Probleme hinsichtlich der Entwicklung des Blattes gehabt.

Neuheuser fügt hinzu, daß es im dokumentarischen Bereich eine DIN-Vorschrift für Erfassungsbelege gebe, woran man sich orientieren könne. So könne der Arbeitsaufwand für den Entwurf verringert werden. In der Archivberatungsstelle verwende man diese Belege mit einigen Änderungen. Man habe nur wenige Wochen gebraucht, um sie zu konzipieren bzw. auf die eigenen Bedürfnisse zurechtzuschneiden. Das habe sich mehr oder weniger zwangsläufig bei der Entwicklung des Systems ergeben. Auf die Frage von Ketelaar wegen der Maske betont er, es handle sich um das Gleiche wie bei einem Formblatt, nur in einer anderen materiellen Form.

Das dezentrale Datenerfassungsgerät, wovon er vorher gesprochen habe, sei nicht mit einem Bildschirm ausgerüstet, sondern arbeite mit Papiausdruck, und auch dabei werde es einen internen Vordruck geben, der vom

System gesteuert werde. Man komme aber immer wieder auf den gleichen Vordruck zurück.

F r e s e kommt auf ein Problem der Klassifizierung zu sprechen. Wenn man einen Bestand verzeichne, ohne daß man gleichzeitig die Klassifikationsmerkmale mit einbringen könne, dürfte es erhebliche Schwierigkeiten geben, die Klassifikation mit unhandlich großen DIN-A4-Formblättern durchzuführen.

R o m e y k stimmt zu, daß die Klassifikationsarbeit mit DIN-A4-Blättern arbeitsintensiver und umständlicher sei als die mit kleinen Verzeichnungsarten. Aber das sei einfach zu umgehen, indem man die eingegebene Verzeichnung auf Karteikärtchen ausdrücke und diese dann mit internen Steuerungszeichen für die Klassifikation versehe, zusätzlich vergebene Prüfnummern würden dafür sorgen, daß die solcher Art ganz konventionell durchgeführte Klassifikation fehlerfrei ausgeworfen werde.

V a n d e V o o r t betont noch einen praktischen Punkt bei der Arbeit am Bildschirm. Seine Erfahrung habe ihm gezeigt, daß es schwierig sei, die Unterlagen und Dokumente durchzusehen und gleichzeitig die Daten am Bildschirm einzubringen. Man sei daher wieder auf zwei getrennte Arbeitsgänge zurückgekommen. (*Dies wurde am Nachmittag bei der Demonstration beim KDC in Nijmegen bestätigt.*) MARDOC sei auf die Formulare zurückgegangen, auch weil die Bearbeiter die Ausarbeitung am Schreibtisch bevorzugten.

K e t e l a r wünscht Informationen über den Stand des EDV-Einsatzes beim Bundesarchiv Koblenz.

H o f m a n berichtet dazu über die Vorgänge im Bundesarchiv. Man verfüge über eine Honeywell-Code-Anlage, die 15 kB speichern kann, 2 Plattenstationen mit je 300 M.B., 3 Bandstationen zu je 120 M.B., 9 Bildschirme, an denen sieben Datentypistinnen arbeiteten. Ferner arbeitete eine Dame an der Datenerfassungsanlage als Operator sowie eine Programmiererin. Die Anlage sei gemietet und koste 160.000 DM im Jahr. Die wichtigste Anwendung liege in der Lager- und Magazinverwaltung. Im Zwischenarchiv in Hanglar sind ungefähr 400.000 – 450.000 Leitzordner mit jeweils 3 bis 4 Vorgängen, also über 1 Million Datensätzen. Auf diesen Akten ruhen einerseits unterschiedliche Aufbewahrungsfristen, andererseits eine 30jährige Sperrfrist für die Benutzung. Am ersten Januar jedes Jahres müsse man wissen, was kassiert werden könne. Dies sei konventionell nicht mehr zu bewältigen, man sei deshalb ungeachtet der Kosten gezwungen, die EDV zu gebrauchen. [Nota: zum Vergleich sei hingewiesen auf die EDV-gesteuerte Verwaltung des Zwischendepots des Archives Nationales in Fontainebleau, der "Cit  des Archives".] Eine andere Anwendung sei der Personalindex. Im Kriege htten viele Leute ihre Rentenunterlagen verloren. Die Bundesanstalt f r Angestelltenversicherung in Berlin, die f r die Berechnung der Rentenanspr che zustndig sei, erkenne auch Anspr che an, wenn der Name einer Person, die einen Anspruch geltend mache, in irgendeiner anderen Akte aufzufinden sei. Jede Akte im Archiv erneut auf einen Namen zu durch-

suchen, sei offensichtlich unmglich. Deswegen habe man alle Akten in einer einzigen groen Aktion ausgewertet und einen Gesamtindex der Namen zusammengestellt. Es handle sich um etwa 800.000 Nennungen. Hier habe der Computer abgesehen von seiner arbeitsersparenden, eine deutlich rechtlich-soziale Funktion.

O p p e l fragt, ob die EDV durch stilistische Analyse eingegebener Dokumente R ckschl sse  ber den Verfasser ermgliche.

A u e r b a c h berichtet  ber die Nachweise von Flschungen bzw. falsche Zuordnung zu einem bestimmten Autor bei alt-russischen Texten. In der Bibliothek von Harvard versuche man dies nachzuweisen, indem man Texte nach ihrem Wortschatz automatisch analysiere. Dieses groe Projekt erfasse mehr als hundert Handschriften.

V a n d e V o o r t f gt noch hinzu, da man in der Zeitschrift "Computer and the Humanities" viel  ber dieses Thema finden knne.

O p p e l fragt, ob es Kontakte gebe zwischen der Archivberatungsstelle Rheinland und dem Hauptstaatsarchiv D sseldorf, um unter Umstnden gleiche Formulare zu entwickeln, die auch anderen Archiven zur Verf gung stehen knnten.

R o m e y k antwortet, da best nden zwei Hauptprobleme. Einerseits gebe es Schwierigkeiten aufgrund der f deralistischen Struktur der Bundesrepublik Deutschland und andererseits auch bei der Frage der Finanzierung, wenn mehrere Partner sich beteiligten. Hier trten insbesondere b rokratische Schwierigkeiten auf.

O p p e l befragt Scholten als Stadtarchivar (mit einem relativ kleinen Dienst), welche Schlufolgerungen er aus dem ziehe, was er in diesen zwei Tagen geh rt habe.

S c h o l t e n meint, da seines Erachtens f r ein kleines Archiv wie Zutphen der Einsatz der EDV nur zu verwirklichen sei in einer Zusammenarbeit von mehreren kleinen Archiven oder in der Kooperation mit einer groeren Institution z.B. einer Universitt.

O p p e l ergnzt, dies setze eine bestimmte Solidaritt voraus, die aber nicht immer vorhanden sei. Dies solle auch als Appell verstanden werden.

K e t e l a r erwhnt, der R ckstand der Niederlndischen Archivverwaltungen sei auch wegen der 50-Jahresfrist mit dem Vorteil verbunden, sich die Erfahrungen anderer Lnder zunutze machen zu knnen. Man suche in den Niederlanden die Zusammenarbeit von staatlichen und kommunalen Archiven, um in einem kleinen Land wie Holland wenigstens ein Projekt anzufangen, das f r Gemeindearchive wie f r die Reichsarchive Lsungen zeitige.

Auch R o e d o e sieht es so, da man in der Bundesrepublik Deutschland weiter bei der Anwendung von Computern sei. Die Schwierigkeiten f r die Archive lgen in:

- der elektronischen Verarbeitung von unstrukturierten Daten und "Images";
- der Wahl der EDV-Systeme;
- den bisweilen sehr großen Archiven und ihrer Lagerverwaltung;
- der Geldbeschaffung;
- der bisweilen zu geringen Größe der Archivverwaltungen, vor allem im kommunalen Bereich.

Mögliche Lösungen seien zu finden in:

- der zwingenden Erhöhung der Dienstleistung und in der erforderlichen Zugänglichkeit. Aufgrund dessen müsse man fraglos in Zukunft auf den Computereinsatz kommen;
- der Zusammenarbeit mit anderen Instituten z.B. Universitäten. Hier falle dem Ministerium vom C.R.M. eine koordinierende Rolle zu. Die ideale Lösung wäre ein zentraler Computer mit verschiedenen kleinen Off-line-Computers (Terminals) in den Archiven. Dies stellte auch die kostengünstigste Lösung dar. Dabei verwies er auf eine Abteilung des Ministeriums, wo an 40 Plätzen mit verschiedenen Systemen gearbeitet worden sei. Die Kosten hätten sich dafür auf 5.2 Millionen Gulden belaufen. Eine Kommission habe festgestellt, daß ein großer zentraler Computer mit 100 Anschlußmöglichkeiten und einem Stab von 10 Personen nur 1.8 Millionen koste.

K e t l a a r schildert abschließend die Situation bezüglich des koordinierten EDV-Einsatzes in den Niederlanden. Die Archive unterstützen dem Kultusministerium (CRM), während die Verwaltung, die Automatisierung und der Einsatz der EDV in der Verwaltung im Ganzen in den Kompetenzbereich des Innenministeriums (BiZa) falle. Jetzt sei eine neue Kommission gegründet worden, die die beiden Bereiche zusammenfasse.

R o e d o e hält es gleichwohl für wichtig, daß von der Basis her Druck ausgeübt werde, damit es nicht nur beim Planen und Reden bleibe.

Im Namen der Veranstalter schloß D r s. M. v a n D r i e l (Rijksarchief in Gelderland) nach dieser Diskussion das Symposium ab, weil die Mehrzahl der Anwesenden an dem am Nachmittag geplanten Besuch des Katholischen Dokumentationszentrums (KDC) an der Katholischen Universität Nijmegen nicht teilnehmen konnte.

Der Sprecher weist darauf hin, daß das Symposium nach Meinung der Veranstalter seinen Zweck erfüllt habe, wenn die Teilnehmer für ihre weitere Arbeit nützliche Erfahrungen gewonnen hätten.

Es sei recht eindrucksvoll gezeigt worden, daß die EDV zur Bewältigung der Aufgaben sowohl im Bereich der Erschließung von Archivgut wie auch in der Archivverwaltung große Hilfeleistungen biete. Zwar bringe die Anwendung der EDV schon ihre eigenen Probleme mit sich und befreie den Archivar nicht gänzlich von seiner konventionellen Erschließungsarbeit, aber sie verspreche doch – wenn auch nach beträchtlichen In-

vestitionen intellektueller und materieller Mittel – dem Archivar und dem Benutzer Neues und Besseres. Allerdings sei auch zu Vorsicht und Mäßigung des Optimismus gemahnt worden. Es sei aber auch mehrfach darauf hingewiesen worden, daß der Archivar den Einsatz der EDV nicht mehr aussparen könne, schon um seinen Aufgaben als künftiger Betreuer EDV-erzeugten Archivguts gewachsen zu sein.

In dieser Hinsicht wurde von allen Teilnehmern eine offenere, erneuerungsfreudigere Haltung der Archivare gefordert. Die jetzige Tagung sei hoffentlich ein Zeichen dieser Aufgeschlossenheit, und trage durch die Berichterstattung dazu bei, sie in weiteren Kreisen zu fördern.

Er schließe den Wunsch an, das Symposium möge beitragen zur Förderung der Zusammenarbeit bei der Entwicklung des archivischen EDV-Einsatzes, der allseits als notwendig gefordert wurde. Es sei auch aufgezeigt worden, daß die EDV nicht eine so esoterische Sache sei, daß sie dem einzelnen Archivar verschlossen bleiben müsse, die Schwierigkeiten nicht unüberwindbar seien, dies um so mehr, wenn man zusammenhalte und sich gegenseitig durch Austausch unterstütze. Diese Kooperation würde Doppelarbeit vermeiden helfen und so der Erneuerung im Archivwesen mehr Nachdruck verleihen. Wenn durch die während dieser zwei Tage entstandenen persönlichen Kontakte sowie durch die Kenntnisnahme der Ergebnisse in weiteren Fachkreisen etwas in dieser Hinsicht erreicht worden sei, hätte das Symposium auch in weiterem Sinne seinen Zweck erfüllt.

Abschließend dankte v a n D r i e l den Teilnehmern für ihre große Beteiligung an den Diskussionen, den Moderatoren für die geleistete Arbeit und den Referenten für ihre Berichte und Erläuterungen, ferner Herrn S o b e c k, dem Betreuer der Technik, der seinen Geburtstag dem Fortschritt der Archivwissenschaft gewidmet habe, Herrn W o u t e r s, dem Protokollführer, dankte er auch namens der Veranstalter. Dank galt auch der Provinz Gelderland für ihre großzügige Gastfreundschaft sowie dem Niederländischen Kultusministerium für die finanzielle Hilfe.

In den Personen von D r. R i c h t e r i n g und D r. O p p e l wurden die Mitveranstalter der Tagung: das Westfälische Archivamt und die Stadt Bocholt, wegen der guten Zusammenarbeit gelobt. Ganz zum Schluß dankte der Sprecher persönlich seinem Kollegen J. H o f m a n, der als eigentlicher Organisator des Symposiums die Hauptlast getragen hat.

Am Nachmittag wurde von einigen Teilnehmern noch das KDC in Nijmegen besucht. Dort hielt D r s. P. S o e t a e r t einen kurzen Vortrag über die Anwendung des RAPIDE-Systems (Repertorium van Automatische Prompte Informatie over de Documentatie-eenheden van het KDC) bei der Archivgut-Erschließung in Zusammenarbeit mit dem Rechenzentrum der Katholischen Universität. Bei diesem Unternehmen sei beabsichtigt, ein integriertes Informationssystem für Archive, Bibliotheken und andere Dokumentationen herzustellen. Abschließend wurde die Eingabe von Daten in den Computer bei der Aufnahme von Aktentiteln demonstriert.

RAPIDE, EIN INFORMATIONSSYSTEM FÜR DIE ERSCHLIESSUNG VERSCHIEDENARTIGER DOKUMENTATION

von P. Soetaert

1. Das KDC in Nimwegen

1.1 Zielsetzung

Im Jahre 1969 wurde von der Katholischen Universität in Nimwegen das Katholiek Documentatie Centrum (KDC) gegründet. Dieses Institut hat zum Ziel:

- eine Dokumentation über das katholische Leben in den Niederlanden vom Anfang des 19. Jahrhunderts bis heute zusammenzustellen;
- diese Dokumentation zu ordnen und zu beschreiben;
- Hilfe beim wissenschaftlichen Studium dieser Dokumentation zu leisten;
- anhand der gesammelten Dokumentation über die katholische Entwicklung in den Niederlanden zu berichten.

1.2 Das Institut

Das KDC gehört zum 'Allgemeinen Dienst' der Katholischen Universität. Zum Stab des Instituts gehören ein Direktor, ein wissenschaftlicher Mitarbeiter, ein Archivar, ein Bibliothekar und eine Sekretärin. Dem wissenschaftlichen Mitarbeiter ist die Koordination der elektronischen Datenverarbeitung im Institut aufgetragen. In ihrer wissenschaftlichen und administrativen Arbeit wird dieser Stab von wechselnden Mitarbeitern und studentischen Assistenten unterstützt.

1.3 Die Dokumentation

Das KDC will dem Benutzer eine möglichst reiche Dokumentation zur Verfügung stellen. Diese Dokumentation wird je nach ihrer Art von einer der drei Abteilungen des KDC betreut:

- dem Archiv (ca. 500 Archive oder Archivalsammlungen von katholischen Organisationen oder Personen);
- der Bibliothek (Bücher, Broschüren und Zeitschriften);
- sonstigen Dokumentationskategorien (u.a. Ausschnitte, Abbildungen, Toten- und Devotionsbildchen, audiovisuelles Material).

2. Das RAPIDE-System

2.1 Der Bedarf einer integrierten Erschließung

Während der ersten zehn Jahre des KDC ist vor allem viel Dokumentationsmaterial gesammelt worden, das nur zum Teil durch Beschreibung erschlossen wurde.

Diese Beschreibung stützt sich möglichst weit auf die im Dokumentationswesen gebräuchlichen verschiedenen Beschreibungstechniken. Dadurch standen dem Benutzer – an und für sich brauchbare – Archivinventare, Bibliothekskataloge und andere Reihen von Beschreibungen zur Verfügung. Parallel dazu wurde es aber immer schwieriger, im Ganzen der KDC-Dokumentation einer für eine bestimmte Untersuchung gewünschten Information auf die Spur zu kommen. Die positiven Erfahrungen bei der Realisierung eines automatisierten Zeitschriftenkatalogs (REPPER) (1), eines Katalogs von etwa 13.000 Zeitschriften und anderer periodischer Druckschriften) war für das KDC eine Herausforderung zu untersuchen, ob und in welchem Maße die Einführung eines automatisierten Informationssystems der globalen Erschließung des in dem KDC gespeicherten Dokumentationsmaterials von Nutzen sein würde.

2.2 Grundgedanke

Aus einem dauernden Gedankenaustausch zwischen Mitarbeitern des KDC und des Rechenzentrums der Nimweger Universität (URC) ging unter der Bezeichnung RAPIDE (2) ein Entwurf für ein integriertes Informationssystem hervor. Ausgangspunkt war das Bedürfnis, mittels eines Suchsystems das gesamte Dokumentationsmaterial nach relevanter Information befragen zu können. Dies bietet dem interessierten Forscher den Vorteil, daß er auf ein und dieselbe Weise in verschiedenen Dokumentations-Kategorien 'suchen' kann. Für die Mitarbeiter der KDC erhöht ein integraler Vorsatz die Möglichkeit, die verschiedenen konventionellen Beschreibungsarten aufeinander abzustimmen. Informationen aus früher beschriebenen Archiven oder Sammlungen stehen nun bei der Herstellung neuer Beschreibungen zur Verfügung.

2.3 Bausteine

Für die Entwicklung eines solchen Informationssystems war es nötig, einige spezifische Wörter zu verwenden. Mit Dokumentationseinheit (3) ist die kleinste Einheit einer als selbständig zu betrachtenden (und also zu beschreibenden) Dokumentation gemeint. Als Dokumentationseinheiten werden demnach betrachtet: ein Buch, eine archivalische Einheit, z.B. ein Aktenbündel, eine Abbildung, eine Reihe von Ausschnitten über denselben Gegenstand usw. Die verschiedenen Elemente der Beschreibung einer Dokumentationseinheit werden Kennzeichen (4) genannt. Jeder Kennzeicheninhalt wird durch einen Kennzeichencode identifiziert. Als Kennzeichen können gelten: der Name

eines Autors, der Titel, die Signatur, die Inventarnummer, die redaktionelle Form, der Inhalt, die Datierung usw. Eine Sammlung solcher Dokumentationseinheiten, die nach den herkömmlichen Beschreibungsweisen als eine zusammenhängende Ganzheit betrachtet werden, heißt eine **K o l l e k t i o n**. In der Archivabteilung entspricht eine RAPIDE-Kollektion einem Bestand. Man wird verstehen, daß es notwendig ist, für jede Kollektion eine spezifische Kombination von Kennzeichen (record format) definieren zu können.

Unter den wichtigsten vorgesehenen Kollektionen können genannt werden: einige hundert Archivbestände und Archivsammlungen, die Bücher und Broschüren, die Zeitschriften, die Ausschnitte und die Abbildungen. Außerdem wird noch erwogen, eine Datenbank aufzubauen, in der Daten über Personen und Organisationen, bibliographische Informationen sowie Daten, die im Rahmen von Forschungsprojekten gesammelt worden sind, gespeichert werden sollen. Sämtliche Kollektionen stellen die RAPIDE **d a t a b a s e** dar.

2.4 Die online-‘Befragung’

Wenn bei der automatisierten Eingabe von Beschreibungen der Dokumentationseinheiten jeder Dokumentationseinheit ein oder mehrere **D e s k r i p t o r e n** zugeordnet werden, ist es nachher möglich, eine Liste von allen für einen am Bildschirm aufgegebenen Deskriptor relevanten Dokumentationseinheiten zu erhalten, ungeachtet, zu welcher Kollektion diese gehören. Als Deskriptor können einer Dokumentationseinheit mitgegeben werden: Themen, Namen von Personen und Organisationen, geographische Namen und chronologische Daten. Es ist auch möglich, bei einer Nachfrage mehrere Deskriptoren einzugeben; diese können dann untereinander durch Operatoren (und, oder, und nicht, von . . . bis . . .) verbunden werden. Als Deskriptoren können nur Wörter/Ausdrücke benutzt werden, welche im **T h e s a u r u s** vorkommen. In diesem Thesaurus sind die gegenseitigen Beziehungen zwischen Deskriptoren, die Verweisungen nach Vorzugsformen usw. festgelegt (5). Neu ist die Absicht, für die Beschreibung sowohl des Archivmaterials und des Bibliotheksmaterials als auch der anderen Dokumentationskategorien einen gemeinsamen Thesaurus zu verwenden. Bei angemessener Überwachung dieses Thesaurus wird gleichzeitig die Informationssprache heranwachsen, die das KDC sowohl für die Beschreibung als auch für das ‘Zurückfinden’ seines Materials braucht.

2.5 Der Printeroutput

Die Verwirklichung von RAPIDE ist nicht nur auf online-Befragung gerichtet. Auch die Produktion von Listen, in denen je Kollektion die Beschreibungen von allen Dokumentationseinheiten ausgedruckt werden, ist möglich. Nicht nur Listen in numerierter oder alphabetischer Folge sind dabei vorgesehen, sondern auch KWIC- und KWOC-Listen (Keyword In bzw. Out Context)

und Indizes anhand der Deskriptoren. Auch die Produktion eines Gesamtindex gehört zu den projektierten Möglichkeiten.

3. Archivanwendungen

Seit September 1981 steht das RAPIDE-System der Archivabteilung des KDC zur Verfügung. Auch wenn im Vergleich mit der Bibliothekautomatisierung die Anwendung des Computers hinsichtlich der Erschließung von Archivmaterial noch nicht so weit entwickelt ist (6), wird die elektronische Datenverarbeitung vorzugsweise für Archivbeschreibungen verwandt, wenn auch vorläufig nur auf der Basis eines breiten Experiments.

3.1 Das record format

Als minimale Kennzeichen für eine Archivbeschreibung gelten in dem KDC:

- die Inventarnummer;
- die redaktionelle Form;
- der Inhalt;
- die Datierung;
- die äußerliche Form.

Eventuell kann jede Beschreibung erweitert werden durch Angaben über:

- das Entwicklungsstadium eines Aktenstücks;
- eine Note (N.N.);
- eine Notierung früherer Codenummern, codeähnlicher Vermerke oder Registraturnummern.

Im Falle einer ‘Sammelbeschreibung’ ist es außerdem notwendig, daß die verschiedenen Einzelteile genannt werden können.

3.2 Die Eingabe

Das Einbringen der Beschreibungen geschieht online anhand eines ‘Frage- und -Antwort-Systems’. Ein **d a t a e n t r y s y s t e m** bittet für jede Dokumentationseinheit darum, ein obligatorisches Kennzeichen auszufüllen. Der am Bildschirm eingebrachte Kennzeicheninhalt wird daraufhin verglichen mit einer für jedes Kennzeichen spezifischen Syntax. Nur gültige Kennzeicheninhalte werden durch RAPIDE akzeptiert und gespeichert. Eine Datierung früher als 1750 oder weitergehend als 1982 wird z.B. nicht akzeptiert. Selbstverständlich ist es auch möglich, nachträgliche Korrekturen durchzuführen oder andere (nicht obligatorische) Kennzeichen hinzuzufügen.

3.3 Die Findbucheinstellung

Im Vergleich zu bereits existierenden Systemen liegt die größte Erneuerung von RAPIDE in der Möglichkeit, den Computer bei der Erstellung von Inventaren einzu-

schalten. Unter Inventar verstehe ich Findbücher nach einem Klassifikationsschema, wie es in der (niederländischen) Archivwelt gebräuchlich ist. Mit Hilfe eines Computers kann man ja die eingegebenen vorläufigen Beschreibungen eines Archivs selektieren und sortieren und sie zu einem Inventar ordnen, worin die Teile des Archivs gruppiert und nach einer durch RAPIDE anerkannten Inventarnummer präsentiert werden. Ich nehme mir vor, über die dabei angewandten Techniken und über die ersten Resultate möglichst schnell zu berichten.

Anmerkungen

1 REPPER = REPertorium van PERiodieken van het KDC, vermeerderd met titels uit het overige bezit van de Katholieke Universiteit Nijmegen.

2 RAPIDE = Repertorium van Automatische Prompte Informatie over de Documentatie-Eenheden van het KDC.

3 Vgl. K.-H. Meyer-Uhlenried: *Methodische Grundlagen für die Planung von Informationssystemen* (München 1977) 13–19.

4 Vgl. den Gebrauch der kenmerkcodes im PICA-Projekt der niederländischen wissenschaftlichen Bibliotheken.

5 Der KDC-Thesaurus wird in den Jahren 1982–1983 ausgearbeitet werden. Hinsichtlich der Formgebung wird dabei *The Urban Information Thesaurus* von W.T. Dürr und P.M. Rosenberg (Westport 1977) als Muster gelten. Vgl. D. Soergel, *Indexing Languages and Thesauri: Construction and Maintenance* (Los Angeles 1974).

6 Einen status quaestionis der Archivautomatisierung in der Welt bietet M. Cook, *Archives and the Computer* (London 1980). Die ausführlichste Literaturübersicht ist: M. Kesner, *Automation, Machine-Readable Records and Archival Administration* (Chicago 1980). Die neueste Literatur wird angezeigt in: ADPA *Automation-Archives-Informatique*, *The American Archivist* und *Der Archivar*.

VERZEICHNIS DER TEILNEHMER

- | | |
|---|--|
| 1. Frau Dr. I. Auerbach | Institut für Archivwissenschaft Marburg |
| 2. J.J.A. Buijlinckx | Gemeentearchief Nijmegen |
| 3. Drs. M. van Driel | Rijksarchief in Gelderland |
| 4. Dr. W. Frese | Westfälisches Archivamt Münster |
| 5. Drs. P.F.A. de Guchteneire | Steinmetzarchief Amsterdam |
| 6. Drs. J. Hofman | Rijksarchief in Gelderland |
| 7. Dr. R. Hofmann | Bundesarchiv Koblenz |
| 8. Prof. Dr. W. Janssen | Hauptstaatsarchiv Düsseldorf |
| 9. Dr. F.C.J. Ketelaar | Centr. Dir. Rijksarchiefdienst Den Haag |
| 10. Mevr. Drs. E. Lievese-Pelser | Gemeentearchief Amsterdam |
| 11. Frau Dr. J. Marschalck | Westfälisches Wirtschaftsarchiv Dortmund |
| 12. Drs. G.J. Mentink | Rijksarchief in Gelderland |
| 13. H.P. Neuheuser | Archivberatungsstelle Rheinland Köln |
| 14. Dr. H.D. Oppel | Stadtarchiv Bocholt |
| 15. Dr. M. Petry | Hauptstaatsarchiv Düsseldorf |
| 16. J. Prins | Provincie Gelderland |
| 17. Dr. H. Richtering | Westfälisches Archivamt Münster |
| 18. R. Riebschläger | Landschaftsverband Rheinland Köln |
| 19. J. Roedoe | Ministerie Volkshuisvesting en Ruimtelijke
Ordering, Den Haag |
| 20. Dr. H. Romeyk | Hauptstaatsarchiv Düsseldorf |
| 21. Drs. C.O.A. Schimmelpenninck van der Oije | Provinciale Archiefinspectie Gelderland |
| 22. Drs. F.W.J. Scholten | Gemeentearchief Zutphen |
| 23. A. Schoonderbeek | Gemeentearchief Arnhem |
| 24. Drs. P. Schotel | Gemeentearchief Dordrecht |
| 25. P. Soertaert | Kath. Documentatiecentrum Nijmegen |
| 26. Drs. J. Spoor | Sansom Automatisering B.V. Alphen a/d Rijn |
| 27. Drs. D. van der Vlis | Stuurgroep Automatisering VAN |
| 28. Dr. J.P. van de Voort | Porject MARDOC, Visserijmuseum Vlaardingen |

VERÖFFENTLICHUNGEN DES LANDSCHAFTSVERBANDES WESTFALEN-LIPPE
WESTFÄLISCHES ARCHIVAMT
Warendorfer Straße 24, 4400 Münster

Inventare nichtstaatlicher Archive Westfalens, Neue Folge (INA NF)

Verlag: Aschendorff Münster; Bezug durch jede Buchhandlung

1. Inventar des Archivs der Stadt Höxter. Bearb. v. Wolfgang Leesch. 1961. XII, 592 S., 18 Abb., kart., 32,— DM
2. Inventar des Grafen v. Spee'schen Archivs Ahausen. Bearb. v. Horst-Oskar Swientek. 1968. XVIII, 825 S., kart., 63,— DM
- 3,1. Inventar des Archivs der Stadt Werl. T. 1: Urkunden. Hrsg. von Rudolf Preisig. 1971. XII, 260 S., 16 Abb., kart., 42,— DM
- 3,2. Inventar des Archivs der Stadt Werl. T.2: Akten. Bearb. v. Dietrich Kausche u. Wolfgang Müller. 1969. XII, 195 S., kart., 22,— DM
4. Inventar des Stadtarchivs Brilon. Bestand A. Bearb. v. Alfred Bruns. 1970. XV, 388 S., 1 Faltkt., kart., 48,— DM
5. Inventar des Fürstlichen Archivs zu Burgsteinfurt. Bestand A. Allgemeine Regierungssachen der Grafschaften Bentheim und Steinfurt. Bearb. v. Alfred Bruns und Wilhelm Kohl, hrsg. von Alfred Bruns. 1971. XIX, 285 S., 1 Faltkt., 89 Abb., davon 7 farbig, kart., 32,— DM
6. Inventar des Fürstlichen Archivs zu Burgsteinfurt. Regierungssachen der Grafschaften Bentheim u. Steinfurt. Bestände A Bentheim, A Steinfurt, G. Bearbeitet von Alfred Bruns und Hans-Joachim Behr. Hrsg. von Alfred Bruns. 1976, XVI und 544 Seiten, 10 Abbildungen auf Tafeln, Leinen, 59,— DM. ISBN 3-402-05278-4
7. Inventar des Fürstlichen Archivs zu Burgsteinfurt. Schulden-, Familien-, Reichs- und Kreisachen. Bestände C, D (Teilbestand), E. Bearbeitet von Alfred Bruns. Erscheint Frühjahr 1982
8. Inventar des Stadtarchivs Kamen. Die Urkunden bis 1500. Bearb. v. Johannes Bauermann. 1978. XX, 188 Seiten, 9 Abbildungen auf Tafeln, Leinen 43,— DM, ISBN 3-402-05282-2

Westfälische Quellen und Archivverzeichnisse (WQA)

Bezug: Westfälisches Archivamt

1. Varlarer Frei- und Wechselbriefe 1329 — 1803. Bearbeitet von Alfred Bruns. 1977. 270 S., 1 Karte, Leinen 19,50 DM
2. Die Abgeordneten des Westfalenparlamentes 1826 — 1978. Bearbeitet von Josef Häming. Eingeleitet und hrsg. von Alfred Bruns. 1978. 702 S., über 1200 Abb., 7 Farbb., Leinen 24,50 DM
3. Inventar des Stadtarchivs Beckum, Bestand A (1238 — 1803). Bearbeitet von Siegfried Schmieder. 1980. 462 S., 43 Abb., 2 Faltkarten, Leinen 28,50 DM
4. Archivverzeichnis Haus Welbergen, Akten. Bearbeitet von Franz Herberhold. 1980. 395 S., 36 Abb., Leinen 27,50 DM
5. Das Archiv des Archidiakonates Billerbeck. Bearbeitet von Alfred Bruns und Peter Löffler. Hrsg. von Alfred Bruns. 1981. 263 S., 50 Abb., 1 Faltkarte, Leinen 25,50 DM
6. Das Archiv des Westfälischen Kunstvereins und des Provinzial-Vereins für Wissenschaft und Kunst. Bearbeitet von Werner Frese. 1981. 246 S., 82 Abb., Leinen 27,00 DM

Nachdrucke zur westfälischen Archivpflege

Bezug: Westfälisches Archivamt

1. Kommunale Registraturordnungen, hrsg. und eingeleitet von Alfred Bruns. 1977. 88 S., broschiert 5,— DM
2. Die Amtssprache. Verdeutschung von Fremdwörtern bei Gerichts- und Verwaltungsbehörden in der Bearbeitung von Karl Bruns. Hrsg. von Alfred Bruns. 2. Aufl. 1980. VIII, 184 S., geheftet 9,— DM
3. Westfalenlexikon 1832 — 1835. Hrsg. und eingeleitet von Alfred Bruns. 1978. XXIV, 293 S., broschiert 19,50 DM

Archivpflege in Westfalen und Lippe

Mitteilungsblatt mit der Beilage „Westfälische Quellen im Bild“, kostenlos hrsg. vom Westfälischen Archivamt

Der westfälische Reichskreis 1708 und 1723. Beilage zur Archivpflege in Westfalen und Lippe 12. 1979, kostenlos.